Supplement to LANGUAGE, Journal of the Linguistic Society
of America

LANGUAGE MONOGRAPHS

PUBLISHED BY THE
LINGUISTIC SOCIETY OF AMERICA

EDITED BY
GEORGE MELVILLE BOLLING
Ohio State University

OLIN H. MOORE
Ohio State University

C. E. PARMENTER
University of Chicago

EDWARD SAPIR
Yale University

NUMBER XVI MARCH, 1935

ON DEFINING THE PHONEME

BY

W. FREEMAN TWADDELL
University of Wisconsin

WAVERLY PRESS, INC.
BALTIMORE

LINGUISTIC SOCIETY OF AMERICA

Founded 1924

OFFICERS FOR 1935

President, Professor Leonard Bloomfield, University of Chicago.

Vice-President, Professor Walter Petersen, University of Chicago.

Secretary and Treasurer, Professor Roland G. Kent, University of Pennsylvania.

Executive Committee, the preceding, and
Professor LeRoy C. Barret, Trinity College.
Professor Clive H. Carruthers, McGill University.
Professor Edward H. Sehrt, George Washington University.

Committee on Publications:
Chairman and Editor: Professor George Melville Bolling, Ohio State University.
To serve through 1935: Professor C. E. Parmenter, University of Chicago.
To serve through 1936: Professor Olin H. Moore, Ohio State University.
To serve through 1937: Professor Edward Sapir, Yale University.

Standing Committee on Research:
Professor E. H. Sturtevant, Chairman, Yale University.
Professor G. M. Bolling, Ohio State University.
Professor Franklin Edgerton, Yale University.

The Linguistic Society of America was founded in December, 1924, for the advancement of the scientific study of language. The Society plans to promote this aim by bringing students of language together in its meetings, and by publishing the fruits of research. It has established a quarterly journal, a series of language monographs, and a series of language dissertations; the last two will appear at irregular intervals, according to the material offered to the Committee on Publications and the funds available for the purpose. Members will receive all in return for the annual dues of Five Dollars.

Membership in the Society is not restricted to professed scholars in linguistics. All persons, whether men or women, who are in sympathy with the objects of the Society, are invited to give it their assistance in furthering its work. Application for membership should be made to the Secretary, Professor Roland G. Kent, University of Pennsylvania, Philadelphia, Pa.

Entered as Second Class Matter at the Postoffice at Baltimore, Maryland.

This Journal is published quarterly by the Linguistic Society of America. Members of the Society receive it without extra charge, three dollars of the annual dues being appropriated for this purpose; to others, its price is five dollars per annum. Subscriptions and other business communications should be addressed to Language, or to Roland G. Kent, Treasurer, L. S. A., University of Pennsylvania, Philadelphia, Pa. Manuscripts for publication should be sent to George Melville Bolling, Ohio State University, Columbus, Ohio.

Supplement to LANGUAGE, Journal of the Linguistic Society
of America

LANGUAGE MONOGRAPHS

PUBLISHED BY THE
LINGUISTIC SOCIETY OF AMERICA

EDITED BY
GEORGE MELVILLE BOLLING
Ohio State University

OLIN H. MOORE
Ohio State University

C. E. PARMENTER
University of Chicago

EDWARD SAPIR
Yale University

NUMBER XVI MARCH, 1935

ON DEFINING THE PHONEME

BY

W. FREEMAN TWADDELL
University of Wisconsin

WAVERLY PRESS, INC.
BALTIMORE

CONTENTS

I. Introduction 5
II. The Phoneme as a mental or psychological reality 8
III. The Phoneme as a physical reality 17
IV. The Phoneme as a fiction 33
V. Definition of the Phoneme as an abstractional, fictitious unit . 37
VI. Comparisons with earlier definitions of the Phoneme 52
VII. Articulatory rather than acoustic characterization of the Phoneme 56
VIII. Some possible applications of an abstractional, fictitious Phoneme 58

I

INTRODUCTION

For many years, now, and particularly within the last decade, there has been increasing use of the term 'phoneme',[1] 'wenn man sich auch nicht immer über seine definition einig ist'.[2] Parallel to this increased use of the term, there has been the re-emergence of a concept of some unit of spoken language, a unit which is not the same as a 'sound of speech'. It has long been known to phoneticians and linguists in general that the sounds of speech, even within the narrowest restrictions of time and place, even within the usage of a single individual, present an almost infinite variety. Scarcely any two speakers of a given dialect pronounce the same word exactly alike, either as to their articulatory movements or as to the sound-waves which those movements set up in the atmosphere. And yet, within the communicative and expressive medium, those (articulatorily and acoustically) slightly different processes are still the same word; and the 'sounds' which comprise this same word are in some way the same sounds, within the frame of that communicative and expressive medium which is the language of the community. It is the recognition of this sameness, this effective unity, which has found expression in the term 'phoneme' as a unit of spoken language.

In many ways, this concept of unity is not recent; its apparent newness is a product of the increased accuracy of phonetic observation which has accumulated evidence of almost infinite variety in the actual 'sounds' of speech (as articulations and sound-waves). It was only after the recognition of actual measurable variety that the recognition of effective structural unity could appear as a contrast, and hence as needing a new technical term: the term which most students of language now use to indicate such an effective structural unit is 'phoneme'.

---

[1] In the preparation of this study I have drawn heavily upon several of my friends. The discussions with Mr. Martin Joos have been especially fruitful in numerous invaluable suggestions and criticisms. My colleagues Professor A. Senn, Professor E. I. Haugen, and Mr. David Sheldon have been most helpful. I am indebted to Professor Leonard Bloomfield for his patience and kindness in discussing the phoneme and our views on it personally and by correspondence.

[2] Jespersen, Linguistica 214.

Before accurate phonetic observation had established the objective variety of the sounds of speech, there was obviously no necessity for such a distinction; earlier students could say 'sound' and mean both the objective articulatory and acoustic phenomenon on the one hand, and the structural unit on the other. Jakob Grimm was in some ways a 'phonologist' (in the sense of the Cercle Linguistique de Prague) precisely because he was not a phonetician; it is regrettable that some contemporary students of language also find it necessary to achieve 'phonological' insight through a resolute disregard of phonetic procedure.

With the increased use of the term 'phoneme' have naturally come attempts by thoughtful linguists to define the term precisely and usefully. It is the purpose of this study to examine the principal definitions of the phoneme, and to propose a new procedure of definition. To some linguists, an examination of fundamental terms and concepts may appear a waste of time. Thus N. S. Trubetzkoy[3]: 'En s'attachant aux formules générales, aux définitions provisoires, qui sont souvent automatiquement répétées même après que l'emploi du terme a changé, on risque de perdre son temps à une critique superficielle et stérile.' And Daniel Jones[4] is 'inclined to think that the terminology usually employed in referring to phonetic phenomena is sufficiently near the mark for all ordinary practical purposes.' But beside this point of view, which is perfectly appropriate for a pioneer, there is certainly justification for a careful examination of the significance of fundamental terms, especially when the significance appears to be far from fixed, in the usage of different linguists. That this is the case with the phoneme will appear in later sections of this study; here it suffices to quote Mr. J. R. Firth[5]: 'The meaning of any ordinary word is subject to change without notice, but technical terms must not be handled in that way. Notice must be given. A word of warning would appear to be necessary with regard to the word "phoneme". What does it mean?' Professor Leonard Bloomfield, whose work exemplifies the value of precise terminology and consistent logic, has briefly stated that value[6]: 'Certain errors can be avoided or corrected by examining and formulating our (at present tacit) assumptions and defining our (often undefined) terms.' Professor Bloomfield has further pointed out[7] the serious consequences of a failure to define clearly and to state assumptions.

---

[3] Journal de Psychologie 30.229.

[4] Le Maître Phonétique (1933) 29.

[5] Le Maître Phonétique (1934) 46; cf. also H. Pedersen in Litteris 5.154.

[6] LANGUAGE 2.153.

[7] LANGUAGE 8.220 ff.; 10.32 ff.; see also Modern Philology 25.211 ff.

The confusion of usage with the term 'phoneme' is unfortunately not unique in linguistic terminology.[8] But in the case of the phoneme this confusion is particularly distressing because of the relatively recent introduction of the term.

To me, at any rate, it has appeared worth while to subject the various definitions and uses of the term 'phoneme' to analytic scrutiny. It has further appeared to me that as between the converse vices of hair-splitting and methodological chaos, the former is preferable.

It will be agreed, I take it, that a satisfactory definition of a fundamental term should meet these criteria: (1) It must be distinctive, i.e. not overlap any other definition of the same order within the frame of study; (2) It must not contain any unstated assumption or ambiguous term; (3) It must be methodologically feasible, i.e. not define an entity which is inaccessible to methods within the frame of study.

The definitions of the phoneme which have been proposed fall into two general groups: definitions in terms of a mental reality, and definitions in terms of a physical reality.[9]

---

[8] Cf. L. Weisgerber, Vorwort zur Probe eines Wörterbuchs der sprachwissenschaftlichen Terminologie 5 ff.; Beiheft IF 51.

[9] The definition to be found in The Cyclopedia of Medicine 6.149 f. (F. A. Davis Company, Philadelphia, 1933) appears to have little bearing on linguistic methodology: 'A *phoneme* is an auditory hallucination consisting of spoken words'.

II

THE PHONEME AS A MENTAL OR PSYCHOLOGICAL REALITY

I have above loosely characterized the difference between sounds and phonemes: 'Within the communicative and expressive medium, those (articulatorily and acoustically) slightly different processes are still the same word, and the "sounds" which comprise this same word are in some way the same sounds.' Another, and even looser, way of formulating the contrast would be this: Although the speakers produce and the hearers experience objectively different sounds, they are not aware of that difference; the speaker intends to produce the same sound, and the hearer has the impression of hearing the same sound. It is in terms of some such formulation that many linguists, particularly those of the school of Baudouin de Courtenay and the members of the Cercle linguistique de Prague and others with related views, have attempted to define the phoneme as a mental reality, as the intention of the speaker or the impression of the hearer, or both.

An anthology of such definitions in terms of mental reality, taken from recent publications, will serve to characterize all essential aspects of this procedure. A. Sommerfelt: Sur la nature du phonème[1]: 'Les phonèmes sont des modèles que les sujets parlants cherchent à imiter' (23). 'L'image articulatoire et auditive d'un phonème donné est toujours le même' (25). 'On ne pourra insister assez sur le caractère psychique des phonèmes' (25).[2] Tytus Benni: Zur neueren Entwicklung des Phonembegriffs[3]: 'Ein Phonem, eine Lautvorstellung, ein gedachter Laut.' N. S. Trubetzkoy: Zur allgemeinen Theorie der phonologischen Vokalsysteme[4]: 'die einzelnen Lautvorstellungen oder Phoneme.' The Cercle linguistique de Prague: Thèses[5]: 'Images acoustico-motrices les plus simples et significatives dans une langue donnée (*phonèmes*).' H. Ułaszyn: Laut, Phonema, Morphonema[6]: 'Diese akustisch-artikulatorischen Vorstellungen sind als solche nicht weiter teilbar. Ich nenne sie "Phonemata"' (54). 'Das Phonema ist also die Vorstellung von einem aus Lautkomplexen abstrahierten selbständigen Laut, der seines semasiologisch-morphologischen Wertes entkleidet ist' (56). 'Für mich ist ein Phonema ein psychisches Äquivalent eines "empirischen Lautes" und zwar eines Lautes, der als Typus empfunden wird, der sich aber unter gewissen Bedingungen ändert' (61).

---

[1] Norsk Tidsskrift for Sprogvidenskap 1.22 ff.

[2] Cf. Trubetzkoy, in Actes du deuxième congrès international de linguistes 109: 'Die Phonologie befasst sich ... mit den Phonemen, d.i. mit den sich in den Sprachlauten realisierenden, im Sprachbewusstsein lebenden Lautabsichten.'

[3] Donum Natalicium Schrijnen (1929) 36.

[4] Travaux du Cercle Linguistique de Prague (hereafter referred to as TCLP) 1.39.

[5] TCLP 1.10 f.

All these definitions agree in the ascription of mental reality to the phoneme, and for me thus fail to meet the requirement of methodological feasibility, i.e. they identify an entity which is inaccessible to scientific methods within the frame of linguistic study. The objections to mentalistic procedure in linguistics have been put by Bloomfield[7] in a form which appears unanswerable and has thus far, at least, been unanswered. It is a work of supererogation to try to restate what Bloomfield has so well stated. For the sake of completeness, though, it may be justifiable to recapitulate the general principles which invalidate any 'mental' definition of the phoneme.

Such a definition is invalid because (1) we have no right to guess about the linguistic workings of an inaccessible 'mind', and (2) we can secure no advantage from such guesses. The linguistic processes of the 'mind' as such are quite simply unobservable; and introspection about linguistic processes is notoriously a fire in a wooden stove. Our only information about the 'mind' is derived from the behavior of the individual whom it inhabits. To interpret that behavior in terms of 'mind' is to commit the logical fallacy of 'explaining' a fact of unknown cause by giving that unknown cause a name, and then citing the named x as the cause of the fact. 'Mind' is indeed a summation of such x's, unknown causes of human behavior.[8]

---

[6] TCLP 4.53 ff.

[7] Modern Philology 25.211 ff.; his Language 32 ff.; and reviews in LANGUAGE 8.220 ff.; 10.32 ff.

[8] The principles presented in this section are limited strictly to applications of scientific methodology, and accordingly do not involve a discussion with those to whom behaviorism is unacceptable as a psychological assumption or (irrelevantly) as a 'philosophy'. Whatever our attitude toward mind, spirit, soul, etc. as realities, we must agree that the scientist proceeds as though there were no such things, as though all his information were acquired through processes of his physiological nervous system. In so far as he occupies himself with psychical, non-material forces, the scientist is not a scientist. The scientific method is quite simply the convention that mind does not exist: science adopts the nominalistic attitude toward the problem of the universals, in matters of procedure.

These logical objections to mentalistic explanation might be suspended (under properly defined assumptions) if there were any pragmatic utility in operating with 'mind' as a linguistic-scientific fiction. But the actual result of mentalistic assumptions is not progress, but obstruction, in scientific work. For the operation with mental 'causes' produces, in Bloomfield's phrase, 'a short-circuit of inquiry'. The procedure of pretending to explain phenomena by giving mentalistic names to their previously unnamed and still unknown causes 'short-circuits inquiry' until some investigator perceives the fraud and returns to the study of phenomena and their correlations. Any correlation of phenomena which can be established on the basis of mental entities or events can also, and more economically, be established on the basis of the phenomena themselves. Such remarks as Trubetzkoy's[9] 'In allen diesen Fällen hat die Akzentbewegung einen Sinn, der auch vom Sprachbewusztsein bis zu einem gewissen Grade empfunden wird' or Baudouin de Courtenay's 'Wörter mit psychisch auslautenden Konsonanten'[10] are meaningless, unless accompanied by phenomenological evidence, and unnecessary if so accompanied.

The application of these principles to the definition of the phoneme is obvious. We can investigate 'mental phonemes' ('psychophones' in the terminology of Tytus Benni[11]) only through our observations of the sounds a speaker produces and the reaction of a hearer to those sounds. Any correlations of those 'mental phonemes' must be established on the basis of sounds produced and reacted to. Any classification in terms of 'intention' or 'impression' is therefore either a needless incursion into the unobservable or an evasion of the task of classifying the sounds as sounds produced and reacted to, or the task of classifying the production of and reaction to the sounds.

E. Sapir is the only 'phonologist', so far as I know, who has perceived the necessity of justifying the ascription of reality to mental sounds and sound-patterns. In his important essay on Sound patterns in language[12], he is aware of 'a possible charge of phonetic metaphysics'. He gives illustrations of the mental reality of certain phonetic patterns. Thus (51) he adduces the difficulty which speakers of English experience in pronouncing initially certain consonants which they pronounce readily in other positions. '*ŋa-* is incredible because there is no *mba-*, *nda-*,

---

[9] Das morphonologische System der russischen Sprache (TCLP 5.2) 35.

[10] Fakultative Sprachlaute, in Donum Natalicium Schrijnen 38.

[11] Donum Natalicium Schrijnen 35.

[12] LANGUAGE 1.37 ff.

ŋ(g)a- series in English.' One may be surprised at finding [ŋ] placed in a series with [mb] and [nd] rather than with the simple nasals [m] and [n]. Sapir has prepared for this surprise; he has said (49): 'in spite of what phoneticians tell us about this sound (*b* : *m* as *d* : *n* as *g* : *ŋ*), no naive English-speaking person can be made to feel in his bones that it belongs to a single series with *m* and *n*. Psychologically it cannot be grouped with them because, unlike them, it is not a freely movable consonant (there are no words beginning with *ŋ*). It still *feels* like *ŋg*, however little it sounds like it.' In other words, [ŋ] is difficult to pronounce initially because it forms a mental series with the initially non-occurrent [mb, nd] rather than with the initially occurrent [m, n]. It belongs with the former group rather than the latter because it does not occur initially. ... In still other words, [ŋ] is difficult to pronounce initially because it does not occur initially.

A rather more valid procedure to suggest the mental reality of a series [mb, nd, ŋ(g)] would be to try to elicit a positive response from a phonetically naive subject. Teach him to pronounce [ŋɑ]; secure from him the statement that it sounds queer; and then ask him to produce some other similar queer sound. When he fails to respond (or responds with [mɑ, nɑ]), suggest [mbɑ], and ask for still another similar queer syllable. If the hypothecated mental series is a reality, he should certainly respond with [ndɑ]. I have made this experiment with several phonetically naive friends, with uniformly negative results.

Sapir has subsequently returned to the problem of demonstrating the mental reality of quasi-phonetic units. His article, La réalité psychologique des phonèmes[13], proceeds on the principle that 'si l'attitude phonémique est plus essentielle, au point de vue psychologique, que l'attitude plus strictement phonétique, on devrait pouvoir la découvrir dans les jugements spontanés énoncés par des sujets qui possèdent complètement leur langue au point de vue pratique, mais qui n'en ont pas une connaissance rationnelle ou consciemment systématique.' To examine in detail each of the incidents which Sapir narrates would take us too far afield; it must suffice here to consider those features which are essential to a determination of the psychological reality of phonemes.

In all these 'experiments', Sapir is obliged to present negative evidence of the real mental existence of a phoneme: i.e., the constant failure of his subjects to record differently sounds which are objectively different. His Paiute guide, in the first case, responds to various kinds of labial consonants with a single symbol [p]. Sapir's procedure is thus to elicit a uniform response to different stimuli. This uniformity of response appears comparable to that of the small child who responds to the sight of various trouser-wearing individuals with a uniform 'Papa'. This response can scarcely be adduced as evidence of the reality of a concept of paternal achievement or potentiality, or of a concept of adult human male, or even a concept of trouser-wearing individual. It might be interpreted as positive evidence of the mental reality of a concept if, to a uniform stimulus, different responses were made. If Sapir's Paiute guide had been trained to record one variety of labial consonant as [p]; if he had been requested to pronounce all other varieties of 'the same sound'; if he had then without hesitation and without duplication responded with examples of all those other varieties—we might then concede that he appeared to have some mental concept of a unified [p]-phoneme which was variously actualized in his speech. But even this demonstration would be a convenience rather than a necessity for linguistic study; it would represent a summary of the behavior of native speakers, a behavior which is already available for the student of language, though in less concentrated form. In each of his experiments, Sapir has to demonstrate the 'correctness' of his guide's uniform responses by a sketch of the pertinent linguistic data, data which he has assembled and interpreted as a trained, accurate, objective linguist, not as an intuitive native speaker.

---

[13] Journal de Psychologie 30.247 ff.

(1) In Southern Paiute, objectively different labial consonants occur postvocalically—fricatives, voiced and voiceless; stops, aspirated and unaspirated, long and short, nasalized and oral. These differences are apparently not primarily or inherently distinctive, but may be regarded as 'largement déterminées par la nature de l'élément, racine ou suffixe, qui les précède, élément qui peut être considéré comme ayant un pouvoir inhérent de spirantisation, de nasalisation ou de gémination' (252). Initially, only a voiceless stop occurs as unnasalized labial. Sapir's guide, whom he describes as 'jeune homme d'intelligence moyenne'[14], surprised him by writing a voiced labial fricative as [p], although objectively of course this [β] differs more from the initial [p] than does the postvocalic long or short stop [p·, p]. We are not told how Tony wrote these latter sounds; if he wrote them, too, as [p], then his procedure is a simple case of uniform response to different stimuli. If Tony wrote the postvocalic stops with a different symbol, then there

---

[14] We learn from the Proceedings of the American Academy of Arts and Sciences 65.2.299, that he was Tony Tillohash, who at that time was 'just about to complete his course at the Carlisle Indian School in Pennsylvania'.

might be some basis for stating ‘que la postvocalique -*β*- est plus étroitement apparentée, fonctionnellement, à la simple initiale *p*- que ne l’est la postvocalique -*p*- (après une voyelle sourde), qui doit toujours être interprétée comme une forme secondaire de -*p*·-.’ (253) But this statement would be strictly inconsistent with Sapir’s report[15]: ‘The tendency to use geminated consonants in composition is probably due to the greater phonetic similarity thus brought about between a simplex and its compound’.

(2) John Whitney, another native interpreter, assured Sapir that the two Sarsi words, *dìní* ‘celui-ci’ and *dìní* ‘cela fait du bruit’, apparently homonyms, ‘étaient totalement différents’. John insisted that the second of these words ended in [t]. As Sapir tells us, both final vowels are aspirated; the increment of aspiration to a vowel may be a positional variant of [t], or it may be a feature of utterance-conclusion. In the one case, John analyzed the aspiration as a [t]-variant, in the other as a finality-feature. The reason for the native speaker’s choice of mode of analysis is, Sapir indicates, the occurrence of morphological variants of *dìní* ‘cela fait du bruit’ which contain an unambiguous [t]. In so far as this incident may be interpreted as evidence of any mental reality, it would appear to be rather a morphological class or lexical unit than any phonetic or quasi-phonetic class or unit.

(3) Similarly, Alex Thomas’s transcription of his native Nootka distinguished between long consonants as positional variants of short consonants and as the product of combination of similar final and initial consonants of separate lexical elements. This practice, too, appears more appropriately referred to lexical than to quasi-phonetic consciousness, in terms of mental reality.

(4) Alex Thomas used the same diacritic modifier to identify the glotticized plosives (plosives followed by glottal stop) and the glotticized nasals (nasals preceded by glottal stop). By disregarding the factor of sequence, according to Sapir, Alex has given us ‘une preuve évidente de la réalité phonologique (*sic*) d’une classe de consonnes glottalisées.’ Whether this procedure is the product of phonetic analysis or of phonological intuition would appear to depend in part, at least, upon the question whether the glottal stop occurs in Nootka as an autonomous distinctive speech-sound or only as an incremental feature of some oral occlusion. If the latter, then Alex’s disregarding the sequential factor is obviously less significant. In my ignorance of Nootka, I can only affirm that the hypothesis of phonetic analysis is

---

[15] Proc. Amer. Acad. 65.1.70.

tenable: Sapir describes Alex as 'l'élève indo-américain le plus brillant que j'aie eu en phonétique pratique'; and phonetic innocence can be rapidly lost.

(5) Sapir's final bit of evidence, the tendency of English-speaking students, taking nonsense-dictation, to indicate an unpronounced glottal stop after short accented final vowels, involves an auditory hallucination of a non-native speech-sound, and is accordingly only negatively and inferentially admissible, at best.

The first and fourth incidents thus represent simple cases of uniform response to different stimuli. The second and third are compound; in these, variants of two different 'phonemes' are objectively very similar (at any rate, elicit a uniform response from the trained phonetician Sapir); the native speaker distinguishes these variants from each other, or better, fails to distinguish each from other forms of its own 'phoneme'. The basis is, as Sapir indicates, a morphological, syntactical, or lexical correlation with other forms. The fifth case involves a positive response, to be sure, but one which must be interpreted negatively with respect to the mental reality which Sapir wishes to demonstrate.

In short, until positive and unambiguous evidence of the mental reality of phonemes can be adduced, it appears methodologically dangerous to define the phoneme in terms of mental reality. Sapir himself defines the phoneme[16] as 'unité qui a une signification fonctionelle dans la forme ou le système rigidement déterminés des sons propres à une langue.' He quotes[17] with apparent approval Bloomfield's non-mentalistic definition, 'a minimum same of vocal feature'. In 1921, Sapir wrote[18]: 'In watching my Nootka interpreter write his language, I often had the curious feeling that he was transcribing an ideal flow of phonetic elements which he heard, inadequately from a purely objective standpoint, as the intention of the actual rumble of speech.' That 'curious feeling' has apparently become a conviction; but it is, I believe, safe to say (and rejoice) that his linguistic methodology has not undergone any comparable metamorphosis. I believe it is fair to say that that methodology does not exemplify, and the unrelated quasi-experimental theorizings do not establish, either the justification or the necessity for making guesses about linguistic mental processes.

It must be emphasized that these criticisms of Sapir's doctrines are not due to any opinion that his position is more vulnerable than that of other mentalistic definers of the phoneme. It is precisely Sapir's acuteness and conscience which appear in his realization of the necessity of justifying the ascription of mental reality to phonemes and phonemic patterns. It is precisely his merit to have avoided the care-free irresponsibility of the others who also consider the phoneme a mental reality, without taking the trouble to give their reasons.[19]

---

[16] Journal de Psychologie 30.247.

[17] In S. A. Rice, Methods in Social Science (1931) 305.

[18] In his Language 58.

There are some definitions which appear to ascribe both mental and physical reality to the phoneme. Thus in W. L. Graff's definition[20]: 'A speech sound represented by a number of unimportant varieties but interpreted as a unit in a given language at a given time', the constituents of a phoneme are physical, but the constitution is mental. The compromise proposal of Otto Jespersen at the Copenhagen Conference belongs here: 'A family of sounds which from an objective point of view may be regarded as distinct, but which are felt naturally by the speakers of a certain language as identical, because they are not used to keep words apart.'

Similarly the definitions of J. Marouzeau[21]: 'Phonème (Laut, Sprechlaut): Élément auquel conduit l'analyse du langage articulé, défini d'une part par la disposition des organes vocaux et les mouvements qui accompagnent ou provoquent le courant d'air expiré ou inspiré, d'autre part par l'impression auditive qui en résulte' (141 f). S.v. Phonologie: '... des phonèmes, c'est-à-dire des sons réalisés dans la langue et propres à exprimer des notions (école linguistique de Prague)' (143).

Finally, there are definitions on the basis of the phoneme's function, i.e. its utilization in distinguishing words and sentences. This utilization is usually, either explicitly or by implication, related to the intention of the speaker. Thus J. Vachek, in What is Phonology?[22], refers to the phoneme as 'a functional unit' and later (83) states: 'The immense, practically unlimited number of sounds and sound variations in every language appears to be a manifestation (or, perhaps, realization) of a definite and strictly limited number of phonemes.' Vilém Mathesius defines phonemes as 'Laute, die in dem analysierten System funktionelle Geltung haben'[23]; as 'sounds endowed with functional values'[24]; and as

---

[19] For a presentation of other grounds of objection to the definition of the phoneme in terms of mental reality, see (TCLP 4): D. Čyževśkyj (3 ff.); K. Bühler (22 ff.); and W. Doroszewski (61 ff.).

[20] In his Language (1932) xl.

[21] Lexique de la terminologie linguistique, Paris, 1933.

[22] English Studies 15.82.

[23] Xenia Pragensia (1929) 435.

[24] Donum Natalicium Schrijnen 47.

'des éléments phonologiques fondamentaux appelés phonèmes, c.-à-d. des sons (simples ou composites) qui ont une valeur fonctionnelle.'[25] And in his article Ziele und Aufgaben der vergleichenden Phonologie[26], he salutes Sapir as one of the founders of 'phonology'. The Cercle linguistique de Prague, which in 1929 defined phonemes as 'images acoustico-motrices'[27], in 1931 shifts the emphasis to a functional characterization: 'Phonème: Unité phonologique non susceptible d'être dissociée en unités phonologiques plus petites et plus simples. (Unité phonologique: Terme d'une opposition phonologique quelconque. Opposition phonologique: Différence phonique susceptible de servir, dans une langue donnée, à la différenciation des significations intellectuelles.)'[28]

[25] TCLP 1.68.
[26] Xenia Pragensia 432 ff.
[27] TCLP 1.10.
[28] TCLP 4.311.

III

THE PHONEME AS A PHYSICAL REALITY

Nearly all students of language who use the term 'phoneme' as in contrast with 'speech-sound' apply the term 'phoneme' to some larger unit. The speech-sounds are, objectively, of an almost infinite variety; the phonemes of a given language are constant. It follows that each larger unit, each phoneme, may correspond to a number of actual speech-sounds. This correspondence may be of two kinds: (1) The phoneme may correspond to some peculiarity or peculiarities, characteristic of all the speech-sounds in question and characteristic of only these speech-sounds; the phoneme is then a feature[1] of the actual speech-sounds. (2) The phoneme may correspond to the sum of all the speech-sounds in question; the phoneme is then a group of speech-sounds. Those who define the phoneme in terms of physical realities follow one of these procedures: Bloomfield, I believe, the first[2]; and Daniel Jones the second[3].

Bloomfield has defined the phoneme on several occasions; each definition requires consideration.

In A set of postulates for the science of language, (LANGUAGE, 2.157) he writes: 'Assumption: Different morphemes may be alike or partly alike as to vocal features. Thus *book* : *table* [b]; *stay* : *west* [st]; *-er* (agent): *-er* (comparative). Definition: A minimum same of vocal feature is a *phoneme* or *distinctive sound.*'

It must first be observed that the assumption ('different morphemes may be alike or partly alike as to vocal features') is itself based upon another assumption of a different order. For we all know that in any occurrence of the events *book* and *table*, those sections which we call [b] are not objectively alike; and this would be even more prominently the case if Bloomfield had found [t] as the "minimum same of vocal feature" in *ten* : *stay* : *writing* : *bottle* : *hat*. He means, of course, that the [b] of *book* and the [b] of *table* are not significantly different within the frame of the American English language. But the equivalence 'not significantly different': 'alike' implies an assumption that some differences of vocal feature are not significant. This assumption Bloomfield formulates (154 f.) as follows: 'Within certain communities successive utterances are alike or partly alike. . . . Linguistics considers only those vocal features which are alike in the two utterances, and only those stimulus-reaction features which are alike in the two utterances. . . . Outside of our science these similarities are only relative; within it they are absolute.'

---

[1] The definition of this term will be discussed later.

[2] The definitions of K. Bühler appear to belong here: '(Jene) lautlichen Besonderheiten, ... die von den einzelnen Sprachen zum Range von diakritischen Zeichen erhoben worden sind' (TCLP 4.44). 'Une invariante qui sert de marque diacritique en vue de distinguer les significations intellectuelles' (TCLP 4.295).

[3] W. Doroszewski's definition may belong here: 'Mais je pense qu'il est plus simple de considérer comme "phonème" le son tel qu'il est réalisé en position isolée, en traitant de "variantes" les produits combinatoires' (TCLP 4.73).

This statement is obviously of prime methodological importance. Bloomfield has defined an utterance as 'an act of speech' and this, in turn, as 'vocal features or sounds'. An utterance, for him, is accordingly an event in a space-time continuum, an event which is correlated to a (potentially) substantially recurrent social situation. Now, the important aspect of successive utterances, as events in a space-time continuum, is that no two of them are alike. It is only linguistically, i.e. as correlated to a recurrent social situation, that two utterances may be alike. Of the practically infinite objective variations between utterances as physical events, not all are linguistically significant. Those physical events are the material of linguistic study, however, and accordingly the first step in linguistic study must be the establishing of criteria for determining which of the innumerable physical differences are also significant linguistic differences. Until those criteria have been established, the use of the terms 'linguistically same' and 'linguistically different' is unjustified. Every one of the succeeding definitions in Bloomfield's Set of postulates, accordingly, appears conditional upon the assumption of the application of correct criteria to determine the linguistic significance of the objective variations in successive utterances. Those criteria can be applied, I believe; otherwise linguistic study is a sheer impossibility. By the failure to expound the criteria whereby linguistic differences are distinguished from physical differences, Bloomfield's definition of the phoneme does not (for me) fulfill the requirements of methodological practicability. It characterizes the difference between a phoneme and a morpheme and a taxeme, etc.; but it is as though a meter were defined as the standard unit of linear measurement. That would suffice to characterize the difference between a meter and a sonnet, or a cumulus cloud; but it would not furnish a basis for application of the unit within the frame of scientific study.

It might be objected that Bloomfield, in his Set of postulates, is merely concerned with setting up a theoretical frame-work of terminology, not with indications of practical procedure. In that case, it would of course be unreasonable to expect his definitions to meet requirements of working technique; but the methodological value of the Set of postulates for the science of language is thereby seriously impaired.

The foregoing observations apply to the logical procedure in Bloomfield's definition of the phoneme in his Set of postulates. In a later definition of the phoneme (in his book Language 77 ff.), he mentions a criterion whereby significant linguistic differences are to be distinguished from the objective physical differences in successive speech-events. The criterion is 'our every-day knowledge', which is a generalization from a large number of experiments. This knowledge can be trusted to 'tell us whether speech-forms are "the same" or "different".' By experimenting with such a speech-form, *pin*, Bloomfield finds certain relationships between this and other forms. Thus: '*pin* ends with the same sound as *fin*, *sin*, *tin*, but begins differently; . . . *pin* begins with the same sound as *pig*, *pill*, *pit*; etc.' 'In this way, we can find forms which partially resemble *pin*, by altering any one of *three* parts of the word . . . we conclude that the distinctive features of this word are three indivisible units. Each of these units occurs also in other combinations, but cannot be further analyzed by partial resemblances: each of the three is *a minimum unit of distinctive sound-feature, a phoneme.*'

Bloomfield summarizes: 'Among the gross acoustic features of any utterance, then, certain ones are distinctive, recurring in recognizable and relatively constant shape in successive utterances. These distinctive features occur in lumps or bundles, each one of which we call a phoneme. The speaker has been trained to make sound-producing movements in such a way that the phoneme-features will be present in the sound-waves, and he has been trained to respond only to these features and to ignore the rest of the gross acoustic mass that reaches his ears. . . . The phonemes of a language are not sounds, but merely features of sound which the speakers have been trained to produce and recognize in the current of actual speech-sound—just as motorists are trained to stop before a red signal, be it an electric signal-light, a lamp, a flag, or what not, although there is no disembodied redness apart from these actual signals.'

In connection with this analogy, it is pertinent to remark that 'although there is no disembodied redness', all these signals do possess a characteristic peculiarity which is constantly associated with their 'redness', viz. the property of emitting or reflecting to the eye of the motorist photo-magnetic impulses in which wave-lengths of about 671 $\mu\mu$ are present. It appears fair to read into Bloomfield's discussion, accordingly, the opinion that a corresponding constant characteristic is a property of every occurrence of a phoneme; this appears to be supported by his reference to 'phoneme-features' as 'present in the sound-waves.'

It must here be noted that Bloomfield makes use of the terms 'feature' and 'features' in what appear to be at least two different senses, none of them explicitly defined. In his Set of postulates (154), he introduces the linguistic discussion thus: 'Psychology, in particular, gives us this series: to certain stimuli (A) a person reacts by speaking; his speech (B) in turn stimulates his hearers to certain reactions (C). By a social habit which every person acquires in infancy from his elders, A-B-C are closely correlated. Within this correlation, the stimuli (A) which cause an act of speech and the reactions (C) which result from it, are very closely linked, because every person acts indifferently as speaker or as hearer. We are free, therefore, without further discussion to speak of *vocal features* or *sounds* (B) and of *stimulus-reaction features* (A-C) of speech. Definition. An act of speech is an *utterance.*' Here the 'vocal features' are equated to 'sounds' (B); they are also 'speech' (B), defined as a sound-continuum which is a reaction to speaker's stimuli (A) and a stimulus to hearer's reactions (C). 'Vocal features' are an interval-fraction of a social situation; 'vocal features' are a sound-continuum, an utterance, a total act of speech. The plurality of the expression 'vocal features' is essentially meaningless, for all we know about them (or it) is the equivalence to 'an act of speech'.

In the same Set of postulates (157), Bloomfield characterizes the phoneme as 'a minimum same of vocal feature'. This shift from plural to singular is bewildering; it implies a transformation of a continuous 'act of speech' into hitherto unmentioned parts, for we have had no clue as to the procedure whereby the continuum 'vocal features' is to be split up into minimum units. All that Bloomfield has previously said about 'vocal features' describes them as an utterance-totality, a continuous event, a coherent, autonomous, and integral fraction of a social event (B in the series A-B-C). Now we are confronted with that continuous event as composed of units; those units are declared to be minimum, and some of them even declared to be 'same'. The term 'vocal features' as synonymous with 'utterance' is potentially confusing; the behind-stage analysis of the otherwise wholly undefined 'vocal features' into a sequence of precisely separated and somehow similar 'minimum sames of vocal feature' definitely is confusing. The mere description of a totality by a plural noun does not justify the claim that by using that noun in the singular the totality is automatically analyzed into its components.

The definition 'minimum same of vocal feature', by this pregnant inflectional shift from 'features' to 'feature' involves, as I see it, two assumptions: (1) that the speech-event has been correctly analyzed into minimum interval-fractions, (2) that some of these minimum fractions are alike in various utterances. The second of these, that 'some morphemes may be alike or partly alike as to vocal features', has already been touched upon. The first assumption is nowhere explicitly stated.

It must be assumed that a substantially similar reasoning is to be followed with the later definition (Language 79): 'a minimum unit of distinctive sound-feature'. It is difficult to associate any meaning with the expression 'unit of . . . feature'; if a 'sound-feature' is understood as a 'sound-continuum', as an event, as a fraction of a social situation, then the expression 'minimum distinctive unit of sound-continuum' may have a meaning.

In the definitions, then, Bloomfield uses a singular form of a term used previously only in the plural, and in its plural use equated to a coherent, autonomous, integral event. That shift involves considerable difficulties.

But new difficulties, involving perhaps another sense of the term 'feature', appear in the course of the discussion of the phoneme in Bloomfield's Language (79 ff.). In that passage, as we have seen, he has defined the phoneme as 'a minimum unit of distinctive sound-feature'. Then he adds: 'The phonemes of a language are not sounds, but merely features of sound which the speakers have been trained to produce and recognize in the current of actual speech-sound. . . . ' A phoneme is thus a 'feature of sound' which the speaker produces. A phoneme must therefore be a part of the sound produced by the speaker. A phoneme must therefore be a fraction of the sound; it must be a fraction of an acoustic event. In so far as acoustic fractions of different utterances are occurrences of the same phoneme, those fractions must be some constant fraction, common to all the utterances. A phoneme thus appears to be some constant fraction, common to all the utterances. Since the 'phoneme-features' are distinctive, a phoneme must be a distinctive or characteristic fraction of each acoustic event. A phoneme thus appears to be a constant, characteristic fraction of phonetic events; a phoneme-feature is then a constant, characteristic acoustic fraction which is present in the sound-waves at each occurrence of that phoneme.[4] Whether a 'phoneme-feature' is a feature *of* a phoneme, or a feature *which is* a phoneme, we are not told: presumably the former. But that Bloomfield appears to accept the existence of some kind of 'phoneme-features', we have already seen.

He describes these features further (Language 79): 'Among the gross acoustic features of any utterance, then, certain ones are distinctive, recurring in recognizable and relatively constant shape in successive utterances.[5] These distinctive features occur in lumps or bundles, each one of which we call a phoneme.' By 'lump or bundle' of features, we are apparently to understand the simultaneous or immediately sequential occurrence of certain acoustic fractions. In so far as a phoneme is a 'lump or bundle' of 'distinctive features', it must be a combination of acoustic fractions (the time correlation not being stated). It is worth noting that if a phoneme is a combination of distinctive features, it can scarcely be defined as 'a minimum same of distinctive sound-feature'.

Despite the difficulty of arriving at any clear formulation of these uses of 'feature', it appears that as Bloomfield defines the phoneme, it is a constant, characteristic fraction or combination of fractions of acoustic events; it is (or is accompanied by) 'phoneme-features in the sound-waves'.

---

[4] These complexities connected with the use of the terms 'feature' and 'features' represent for me the greatest difficulty in determining precisely the meaning of Bloomfield's definitions of the phoneme. That difficulty is due not only to the lack of any preliminary definition of the terms as applied to events, but also to the unannounced introduction of different usages for the term. At all events, it is hard to see how any definition of 'phoneme' or 'feature' could warrant the optimism with which Bloomfield writes (Modern Philology 25.217): 'The logical demand that a science speak in quantitative terms is met by linguistics because it speaks in terms of phonemes'. The phoneme as a quantitative unit (even of feature) is a difficult enough concept. But these units are by no means quantitatively comparable; one [r]-phoneme is not equal to 3.76 [p]-phonemes; and unless quantitative units are quantitatively comparable, it appears unjustified to claim that we 'speak in quantitative terms'.

[5] In so far as the 'constant shape in successive utterances' is only 'relative', the recurrence of the phonemes is only relative. But if the phoneme is 'a minimum same'; if 'the distinctive features of the word (*pin*) are three indivisible units', and 'each of these units occurs also in other combinations' (Language 79), then the recurrence must in some sense be absolute; and in that same sense the 'shape in successive utterances' must be absolutely constant. If that 'shape' is only relatively constant, then the recurrence is only relative, the sameness of phonemes in successive utterances is only relative; and the entire procedure is necessarily invalid.

If such features existed, the determination of the phonemes of a given language, and the definition of the phoneme, would be achieved. We should ask Bloomfield to write out a text in his admirably practicable phonemic transcription and read it into a high-quality oscillograph hook up. We should then examine the record of his total acoustic production. We should, to be sure, find much in the 'gross acoustic mass' which was non-distinctive; but we should also find the 'phoneme-features'. On every occasion which corresponded to a [t] in his transcription, we should find in the acoustic record some constant, characteristic fraction of sound-waves.

That we do not find any such constant, characteristic fraction is of course a commonplace of experimental phonetics. If Bloomfield is right, the only explanation for our failure to find 'phoneme-features' in an acoustic record is the inadequacy of our recording apparatus. Such appears to be the explanation Bloomfield would advance; in an outline of an address on The Phoneme to the Language and Literature Club of the University of Wisconsin, 13 April 1934, he states: 'The physical (acoustic) definition of each phoneme of any given dialect can be expected to come from the laboratory within the next decades.'

Two queries suggest themselves here. The first is the question as to the propriety of basing a definition (for it is the recurrence of the 'phoneme-features' which Bloomfield adduces as evidence of the sameness of these minimum units) upon a discovery which is envisaged only as a future possibility. The more serious question is this: What is the basis for the belief that future laboratory research will possess more adequate apparatus? Granted that our present acoustic records are far from perfect; granted that considerable technical improvements are to be expected: it remains, however, extremely doubtful whether these improvements will be of a sort to reveal what Bloomfield hopes to find, a set of records which will enable an observer to determine a constant, characteristic fraction of sound-wave, common to all occurrences of 'the same phoneme'. For whatever that (hypothecated) fraction may be, we have no reason to believe that it is beyond the range of our present recording apparatus.

The researches of C. Stumpf, reported in the Handbuch der Physik (Berlin 1927) 8.456, indicate that with an upper transmission limit of 6020 cycles, the only acoustic disturbance was a slight 'blunting' of [s]

Even with so low an upper limit as 2461 cycles, the act of speech is reported as being 'noch ganz gut verständlich', though certain distortions of front vowels and voiceless fricatives are noted. It appears justifiable, then, to assert that the frequencies up to 6500 cycles contain all the distinctive acoustic 'features' of speech. As a matter of fact, many phonograph records and radio loud-speakers transmit speech with a much lower limit (not to speak of the telephone); yet the distinctive sound-features must be present, for it is a matter of 'our every-day knowledge' that we 'understand' that speech.

But our present recording apparatus has a fairly uniform response up to at least 8000 cycles[6]. All those 'phoneme-features' which are associated with pitch and timbre (frequency and wave-form) must therefore be present in our records. The factors of duration and intensity (amplitude) are of course very precisely determinable.

The laboratory situation appears, then, to be as follows: We know the range within which the distinctive features must lie; we have records covering that range. Yet those records do not reveal any constant characteristic of sound-waves (an acoustic feature) for each phoneme; even with the acoustically relatively simple vowels the so-called 'formants' are only proximately determinable, and within very wide ranges.

The only legitimate conclusion at present appears to be that the presence of 'phoneme-features' as positive, additive entities in the sound-waves is not demonstrable, and there is no reason to believe that it will be.

It has been necessary to consider at some length this question of the existence of 'phoneme-features in the sound-waves', for this existence appears to be an indispensable part of Bloomfield's definition. He is justified in calling various sections of utterance-events occurrences of the *same* phoneme only if he can establish some recurrent fraction, some sameness in all such occurrences. By the discovery that '*pin* ends with the same sound as *fin, sin, tin,* but begins differently', he is justified in saying that the [p] of *pin* is a different phoneme from the [f, s, t] of *fin, sin, tin.* But before he may 'recognize a phoneme even when it appears in different parts of words, as *pin, apple, mop*', he will need more than 'a little practice'. He must establish some constant characteristic fraction.

By setting up the groups (1) *pin* : *fin, sin, tin*; (2) *apple* : *addle*; (3) *mop* : *moll,* we obtain the following relations: (1) [p] which is different from [f, s, t]; (2) [p] which is different from [d]; (3) [p] which is different from [l]. But the grouping together of these various [p]s is legitimate only when there is positive evidence of their constituting a positive entity. As K. Jimbo, quite soundly, says[7]: 'We cannot assert from this without dogmatism that the [a] of *ame* and the [a] of *aka* are exactly the same. They *may* be the same, but we cannot prove the identity until due examination, investigation and experiment be made.' The evidence of identity which Bloomfield appears to regard as appropriate is the presence of 'phoneme-features in the sound-waves'. That evidence cannot yet be adduced, and the expectation that it will be forthcoming is itself in need of justification.[8]

---

[6] Cf. Hdb. d. Phys. 8.559 ff., 603; H. Fletcher, Speech and Hearing 27.

Daniel Jones has given definitions of the phoneme in three significant passages: (1) in Trofimov and Jones, The Pronunciation of Russian (1923) 49 ff.; (2) in his notes On Phonemes (TCLP 4.74); (3) in the course of the proceedings of the International Congress of Phonetic Sciences at Amsterdam, 1932. The definitions are: (1) 'A phoneme is a group of sounds consisting of an important sound of the language (i.e. the most frequently used member of that group) together with others which take its place in particular sound-groups.' (2) 'A Phoneme may be defined as a family of sounds in a given language, consisting of an important sound of the language together with other related sounds, which take its place in particular sound-sequences.' (3) 'Definition of a phoneme: a family of sounds in a given language, which are related in character and are such that no one of them ever occurs in the same surroundings as any other in words.'

It will be observed that these definitions represent a progress toward precision. The elimination of the ambiguous uses of the term 'group' in the first definition is of course desirable. The addition of the term 'related' in the second and third definitions appears to make difficult the connection of the English sounds [h] and [ŋ] in a single phoneme; under the first formulation, such a combination would be possible: For [h] occurs only before stressed vowels and [w, j], and might thus be said to 'take the place' of [ŋ], which never occurs in these positions. Presumably the use of the somewhat fanciful term 'family' in the second and third definitions is to be understood as emphasizing the necessity of an organogenetic relationship of the various sounds which are to be grouped together in a phoneme. The second and third definitions differ in the elimination of the distinction between 'important' and 'unimportant' members of a phoneme.

---

[7] In H. E. Palmer, The Principles of Romanization (Tokyo 1931) 44.

[8] Further, quite apart from technical considerations, it is difficult to see what reason there should be for expecting to find any common acoustic characteristic of the [p]s and [t]s, respectively, in *pit* and *tip*, especially if the final stops are pronounced without explosion.

Essentially, it is fair to say, the three definitions agree. A phoneme is for Jones a 'group or family of sounds'. It is clear from his use of terms that by 'sound' he understands an abstraction of a large number of separate sound-events differing so little from one another that the difference is imperceptible. The different sounds which compose a phoneme are related sounds; i.e. they possess or are marked by similarities in the organogenetic processes which produce them. Finally, these different but organogenetically similar sounds alternate with one another in an absolute correlation to particular (preceding and succeeding) sound-sequences.

Sounds which alternate otherwise than in correlation to particular sound-sequences are called 'variphones' by Jones. 'It occasionally happens that a speaker uses one of two or more sounds absolutely indifferently and apparently at random. The Japanese *r* furnishes a good illustration. A single speaker will sometimes pronounce it very much like an English *r* and sometimes as a sound resembling *l*. But he does not do this according to a definite system as in the case of members of the same phoneme; when he says any particular word (say *miru* 'to see') he sometimes says it with English *r* and sometimes with this *l*-like sound and sometimes with an actual *l*. He is unaware that his pronunciation varies. These three sounds may be said to constitute a variphone.'[9] In Jones's terminology, the expressions 'sound, phoneme, variphone' appear to correspond respectively to 'phone, contactual phoneme, free phoneme' in the terminology which H. E. Palmer had made use of in The Principles of Romanization.[10]

Both the phoneme and the variphone, then, are groups of related sounds which alternate, the members of the phoneme systematically, the members of the variphone unsystematically. The organization of these different sounds into a group depends upon their being related; the organization within the group depends upon the nature of their alternation. It is the organization into groups which concerns us here; therein the Jones phoneme and variphone need not be distinguished.

A phoneme or a variphone is a group of sounds which are related, i.e.

---

[9] In Proceedings of the International Congress of Phonetic Sciences, reprint distributed with Le Maître Phonétique (1933).

[10] It should be noted that Palmer explicitly describes the 'phone' as an abstraction.

which are, as a group, organogenetically similar to each other, and, as a group, organogenetically different from the sounds of all other groups within the language.[11] There is a minimum of organogenetic similarity without which two sounds cannot be considered members of the same phoneme or variphone. There must, accordingly, be a hiatus of similarity between the members of one group and the members of another. For 'the same sound' obviously cannot be a member of two different phonemes or variphones, since it would then be more similar to itself than to the other members of its respective phonemes or variphones. Thus, if a sound x could belong to both phonemes A and B, then Ax would have greater similarity to one member of the B phoneme (namely Bx) than to the other members of the A phoneme. Further, such a various distribution of the 'same sound' would be possible only on the basis of a principle of autonomy on the part of the phoneme or variphone itself; since those units are described as 'groups of sounds', the grouping must be in terms of the nature of the sounds themselves.

It appears that Jones's task in determining a phoneme is in a way the converse of Bloomfield's. Whereas Bloomfield has to determine some constant characteristic (acoustic) fraction which marks all members of the group, Jones has to determine the hiatus of organogenetic similarity which sets off all members of the group as distinguished from the members of all other groups.

The difficulty of establishing such a hiatus of similarity between the sounds of two groups can be indicated by an example from the usage of many speakers of General American English. If we examine the vowels of the two lists of words (1) *den, ell, mess, bet* and (2) *Dane, ale, mace, bait*, it is not hard to find that the vowels of each list do constitute such groups of related, organogenetically similar sounds. We can establish a hiatus of similarity between the members of the two groups. But when we consider the vowels of such words as *dare, air, mare, bear*, we discover a sound which lies squarely across our hiatus. The situation is complicated by our failure to find in any monosyllable ending with *-r* a vowel which falls entirely within the similarity-boundaries previously set up for either of our two phonemes.[12] To assign the sound of *dare*, *air*, etc. to one or the other phoneme appears arbitrary[13]; to assign it to both is illogical; to assign it to neither (by calling it, in connection with the -*r*, a diphthong) is to set a precedent which will involve the creation of a large number of such phonemes, and thus disturb the numerical strict limitation which most definers of the phoneme impose. Further, such a procedure would lead to difficulty in classifying the vowels of many a speaker of General American, who sometimes pronounces *Mary* and *merry* as homophones, with the sound of *mare*, at other times pronounces them differently. (We should then have two variphones, the constituent members of which would be organogenetically related to the members of two of three different phonemes!)

---

[11] Cf. Definition of a phoneme (Le Maître Phonétique [1929] 43 f.): 'Two sounds of a language are said to belong to the same phoneme when they are related in character and it so happens that one of them never occurs in exactly the same situation as the other in connected sentences of that language. ... It would indeed be possible to group the sounds of a language into phonemes without knowing the meanings of any words.' Query: without knowing that they *are* words and different words? Cf. Jones's discussion (Le Maître Phonétique [1933] 30) of Palmer's principle of 'multiple identity'.

To be sure, consistent and comprehensible transcriptions of General American can be written by assigning the vowel of *dare*, *air*, etc., to either of the phonemes in question. Thus Bloomfield assigns it to the same phoneme as *Dane*, *ale*; C. K. Thomas, for example, assigns it to the same phoneme as *den*, *ell*. These two transcribers are not necessarily indicating a difference of phoneme usage in their speech; they may be simply adopting different conventions. That both conventions are transcriptionally practicable does not alter the arbitrary nature of either.

If Jones's definition of the phoneme (and variphone) appears inadequate as a theoretical basis for the study of phonetic relations within a language, it is probably because he had no such aim in view. He would probably agree with J. Vachek[14]: 'The phoneme of Professor Jones is a unit the establishing of which was due to purposes of phonetic transcription, not to the study of language as a functioning system.' Certainly Jones has always professed himself satisfied with a definition of a unit for purposes of transcription. In his paper On Phonemes (TCLP 4.78) he writes: 'The main object of grouping the sounds of a language together into phonemes is to establish a simple and adequate way of writing the language.' And at Amsterdam in 1932, under the heading 'Practical importance of the phoneme', he observed: 'The grouping of sounds into phonemes enables us to construct the simplest systems of phonetic transcription for every language.'

The purposes of transcription and of the study of phonetic relations within a language are not the same. It is not surprising that a unit defined for the one purpose should not be wholly applicable to the other. After all, a perfectly consistent and unambiguous transcription of English is possible with the use of a single symbol in the place of [h] and [ŋ]; but no phonetician or phonologist could consent to regarding these two 'sounds' as variants of the same phoneme.

---

[12] Cf. Bloomfield, Language 124 ff.

[13] Particularly since it could fit equally well into the 'internal organization' within either group.

[14] English Studies 15.89. Cf. also Vachek's Daniel Jones and the Phoneme, in Charisteria Guilelmo Mathesio (Prague 1932) 25 ff.

The units with which Harold E. Palmer operates, in The Principles of Romanization, appear to be basically similar to those of Jones. Palmer's generic term for 'group of sounds', the *metaphone*, is to be sure defined as 'two or more phones which serve jointly as units of meaning within the limits of a given linguistic community' (thereby giving rise to cases of 'multiple identity'); and the introduction of considerations of meaning is quite properly avoided by Jones. But the various kinds of metaphones appear to be legitimately equated with various of the Jones units. Thus, the Palmer contactual phoneme corresponds to the Jones phoneme, and the Palmer free phoneme to the Jones variphone. What Palmer calls *dynamophones* (e.g., the phones or 'sounds' of the vowel of *at* in 'He wasn't at home' and 'He's always at it'), may be properly regarded as contactual phonemes, if we follow Bloomfield in admitting stress as a phoneme. (The only objection raised by Jones[15] to Bloomfield's procedure in considering stress a phoneme is on the basis of the traditional use of the term 'phoneme'; it should be clear by now that the tradition is scarcely consistent.) The two other kinds of metaphones, the *diaphone* and the *phonogene*, involve more than one dialect or more than one chronological stage, and accordingly fall outside the scope of this study.

The units of Palmer thus appear to be strictly comparable to those of Jones, and the observations above therefore equally applicable to them. It is fair to note that although Palmer proceeds with much more cautious and explicit logic, his purpose is essentially the same as Jones's: the establishing of sanctions for the graphical representation of a language.

It appears proper to include this mention of Palmer's units in a section dealing with definitions of the phoneme in terms of physical reality, despite his rather frequent references to 'Lautvorstellungen', 'the feeling that native language-users have towards the phonetic units', etc. For he quotes[16] with apparent approval Jimbo's note: 'The treating of "abstract" sounds does not mean the treating of empty, psychical, non-existing, chimerical notions or ideas. What we are treating are concrete sounds, but only those parts of concrete sounds that have been "abstracted" or "picked out".'

---

[15] Le Maître Phonétique (1933) 53; (1934) 54.

[16] Page 45, footnote 2.

The most comprehensive and methodologically lucid treatment of the phoneme with which I am acquainted is The phonemic principle, by Morris Swadesh.[17] Swadesh has recognized and avoided many of the palpable weaknesses of earlier writings on the subject. The dangers of a treatment which is at once comprehensive and concise are eclecticism and circularity; and it is possible that Swadesh's treatment is not invulnerable on these counts. Thus, a phoneme is defined as one of 'a limited number of elemental types of speech sounds' (the term 'type' is not defined); and yet 'the phonemes of a language are, in a sense, percepts to the native speakers of the given language'. Can 'type' be so defined that 'a type of speech sound' is in any sense 'a percept to the native speakers'? And it appears circular to apply 'pattern congruity', which is certainly an attribute of the group-totality of phonemes, as a criterion in the determination of any particular phoneme.

But these objections would be minor and probably relate more to details of phrasing than to any fundamental procedures. A much more important objection is that Swadesh, too, leaves open the possibility of arbitrary procedure in the application of his method. One of his tests for the discovery of the phonemes of a language is (123): '*The criterion of complementary distribution.* If it is true of two similar types of sounds that only one of them normally occurs in certain phonetic surroundings and that only the other normally occurs in certain other phonetic surroundings, the two may be sub-types of the same phoneme. If the distribution of one type of sound is complementary to that of more than one other, it is to be identified with one rather than the other if there is a more definite phonetic similarity in that direction; an example is the *p* of English *speech* whose distribution is complementary to that of the voiced labial *b* as well as to that of the voiceless labial stop sounds of *peak*, *keep*, *happen*, but goes with the latter rather than the former because of the phonetic similarity.' This case may not be so troublesome as the disposition of, say, the vowel in *sing*, but it may be taken as an example of what Swadesh calls multiple 'complementary distribution'.

Swadesh has said (119): 'The norm of the phoneme may be a multiple one. That is, instead of one norm, there may be two or more. Such variant norms are ordinarily conditional, depending on the phonetic surroundings in which the phoneme occurs. Thus one may distinguish at least three norms for English *p*:

---

[17] LANGUAGE 10.117 ff.

(1) Relatively fortis, aspirated: e.g., in initial position, as in *pit.*
(2) Fortis unaspirated: e.g., medially between vowels, as in *upper.*
(3) Lenis unaspirated: e.g., after *s*, as in *spill.*'

Let us consider, on the basis of these descriptions, the bilabial stops in the words *pill, upper, spill, bill.*

1. *pill:* fortis, aspirated, voiceless.
2. *upper:* fortis, unaspirated, voiceless.
3. *spill:* lenis, unaspirated, voiceless.
4. *bill:* lenis, unaspirated, voiced.

The first and fourth sound-types obviously represent the maximal differentiation of the two phonemes. In the position in which the two phonemes are most decisively differentiated, *p* is fortis, aspirated, voiceless; and *b* is lenis, unaspirated, voiced. Taking, for a moment, those factors as characteristic, we may describe our four sounds thus:

1. *pill:* tension, p; aspiration, p; voice, p. (p p p)
2. *upper:* tension, p; aspiration, b; voice, p. (p b p)
3. *spill:* tension, b; aspiration, b; voice, p. (b b p)
4. *bill:* tension, b; aspiration, b; voice, b. (b b b)

The third sound, which is in complementary distribution to each of the others, differs from the second and fourth in one articulatory factor, from the first in two. What, except the conventional spelling, justifies the inclusion of the stop of *spill* in a [p]-phoneme? Is it any more arbitrary to regard this stop as a conditional variant of [b] (voiceless after a voiceless fricative)? It must further be remembered that the arbitrary ascription of any decisive potency to the factor of voice will lead to insoluble difficulties in the disposition of the 'voiced *t*' in *latter* and *writing* as compared with *ladder* and *riding*, in General American.

It appears arbitrary, in the absence of more decisive 'phonetic similarity', to assign the stop of *spill* to [p] *or to* [b]. Arbitrary, from the point of view of the study of phonetic relations within the language, where a [b]-phoneme without a post-sibilant member is just as defective as a similarly defective [p]-phoneme; arbitrary, even, from the relatively trivial point of view of phonetic transcription, for an equally consistent and unambiguous transcription could be written with *sb-* instead of *sp-* in such words.

The inevitability of such an arbitrary procedure in phoneme-determination could be removed only by that discovery of positive phoneme-features which Bloomfield expects. The dilemma appears to be absolute, if the phoneme is defined in terms of observable phonemena. In cases of multiple complementary distribution, a 'sound' must be assigned to one of several phonemes arbitrarily. The only way out is the hypothesis of a positive observable phoneme-feature; Bloomfield, in electing this procedure, achieves coherent structure in his treatment of the phoneme, but must accept the onus of discovering and determining those features. Pending that discovery and determination, the numerous instances of multiple complementary distribution will furnish difficulties which can often be solved only on grounds of arbitrary preference or convention—especially since it is very often precisely in such cases of multiple complementary distribution that the speakers of a language grieve us by their malignant perversity in producing sounds which are not unambiguous as to ‘phonetic similarity’.

IV

THE PHONEME AS A FICTION

All attempts to associate the term 'phoneme' with either a mental or a physical reality appear open to serious if not unanswerable objection. And although the wide and increasing use of the term by linguistic students indicates that it is an eminently useful term (for that use cannot be explained by any pleasure in using esoteric or complicated terminology), it is obviously undesirable to have that term questionably defined. The stature of the scholars who have presented definitions is such that it would be difficult if not impertinent to hope that any one else could avoid the difficulties they encountered. Where Bloomfield, Sapir, Daniel Jones, Trubetzkoy, and the others have not been completely successful, it appears improbable that their methods of definition could lead to better results in other hands.

Is a different procedure possible? Although these two procedures of definition for the phoneme—regarding it as a physical reality of some order or as a mental (or psychological) reality—appear to represent the two possibilities, perhaps they are only subalternatives of one of two possible procedures. It may be that before it is appropriate to answer or even to ask the question as to the nature of the reality associated with the term 'phoneme', we should first ask and answer the question whether the term can be profitably associated with any reality at all. It is what might be called the thesis of this paper that it is inexpedient and probably impossible (at present) to associate the term with a reality: probably impossible, because the attempts by competent and conscientious linguists to define the phoneme in terms of reality have not been wholly satisfactory; inexpedient, because the purposes to which the term may be put in our discipline are served equally well or better by regarding the phoneme as an abstractional, fictitious unit.

If a satisfactory procedure of definition of the phoneme as an abstractional fiction can be worked out, that definition would have two advantages: (1) we could use the phoneme as a terminological convenience just as well as at present; (2) we should not have to use as legal currency in our study the promissory notes of the laboratory, which are liable to a heavy discount. Then, if the next decades should bring a satisfactory physical definition of the phoneme from the laboratory, we could discard the abstractional and substitute the physical definition, without disturbing the products of earlier linguistic study. If advances in psychological methods, at present unimaginable, should justify the use of psychological units in linguistic study, we could substitute a psychological definition without the necessity of reappraising all preceding phonological research. We could simply say: 'The phoneme, which we have been cautiously regarding as a convenient fiction, appears to be an objectively determinable reality. How nice!' In the happy event that the phoneme should prove determinable both physically and psychologically, the abstractional fiction definition might prove the method of rapprochement between the two determinations, a rapprochement of significance far beyond the field of linguistics. But pending those physical or psychological determinations, we should be able to use the term 'phoneme' without being subjected to the embarassment of having in our discipline a fundamental unit which is undetermined and the nature of which is a matter of wide disagreement among linguists.

The reluctance to attribute reality or even linguistic autonomy to the phoneme is nothing new. Otto Jespersen could with comfort reprint in 1933[1] some remarks he had made on this subject almost half a century earlier. He summarizes the basic Junggrammatiker postulate: 'denn das organgefühl wird ja nicht für das einzelne wort, sondern für den einzelnen laut gebildet; die aussprache wird nicht für jedes einzelne wort gelernt.' He goes on to say: 'Gegen diese letzte kategorische behauptung wäre man versucht die gerade entgegengesetzte zu stellen: die aussprache muss immer für jedes einzelne wort gelernt werden, denn die aussprache ist ja nichts als das wort selbst, oder besser die eine, die nach aussen gekehrte seite des wortes im gegensatz zu der innern der bedeutung.' Jespersen does not of course maintain this extreme statement in all literalness; but as between the two formulations of the relation of 'sound' (or phoneme) and 'word' (or free form), his is incomparably the more clear-headed and linguistically valid.

Bloomfield goes at least part of the way with Jespersen: he observes[2]: 'The phoneme is an abstraction obtained from series of utterances, e.g., the English phoneme [t] from [tin, win, pin], etc., and [tin, tuk, it], etc. An ear trained to other languages will hear differences between the [t]'s of [tik, stik, botr, bit] which are not distinctive, that is, in English non-linguistic.' This characterization of the phoneme as an abstraction is not necessarily inconsistent, of course, with Bloomfield's definitions of the phoneme in terms of acoustic reality. He means, I take it, that the determination of a phoneme is a process of abstracting from the utterance (or 'word') as a gross acoustic event certain constant characteristic phoneme-features and disregarding the variable non-distinctive features. But when he uses (217) such a formula as 'This word consists of such and such of the basic phonemes of my language arranged in such and such sequence', it is apparent that he is following essentially the Junggrammatiker procedure and regards the phonemes as positive additive entities.

---

[1] Linguistica 180 f.

[2] Modern Philology 25.216 f.

The theoretical objections to regarding the phonemes as positive additive entities, advanced by Jespersen in terms of learning, can be supported by the evidence of actual linguistic behavior. We respond differently to 'pen' and 'pin'. But it cannot be demonstrated that this difference is conditioned by any psychic isolation of the second phoneme in each word; it is even questionable whether the response to the isolated acoustic complexes [ϵ] and [ɪ] may properly be called linguistic. If I ask for a pen and you bring me a pin, I correct you by saying 'No; I wanted a *pen*'. Only if I were speaking to another student of phonetics—that is, in an abnormal social situation—would I say anything like '[ϵ], not [ɪ]'; or 'the half-open front vowel, not the close lax one'. Even when longer words are involved (say *illusion* and *allusion*), the correction is made not by isolating the 'misheard phoneme' but by emphasizing it in its phonetic context through an unconventional incidence of accent.

On the side of the perception of spoken language, a similar irrelevance of the existential reality of phonemes as such is a commonplace. Cf. C. Stumpf, Die Sprachlaute (1926), 94 f.: 'Auszerordentlich auffallend ist bei solchen Versuchen in jedem Stadium der Zerstörung, wie lange doch noch die zusammenhängende sinnvolle Rede verständlich bleibt, nachdem sehr maszgebende Bestandteile für sich allein bereits stark alteriert, ja unkenntlich geworden sind. Man kann sich diesen, im allgemeinen ja bekannten Einflusz des Zusammenhanges auf die richtige Auffassung gar nicht schlagender als durch solche Versuche vergegenwärtigen und ist immer wieder davon überrascht. Bei der Einstellung 6, 7[3] wurde selbst ein Wort wie "still" noch erkannt, obgleich so gut wie alle Bestandteile einzeln unkenntlich waren. Ebenso noch bei Einstel-

---

[3] Upper limit of 1230 cycles.

lung 9, 6[4] "insbesondere", wo höchstens o und r für sich erkennbar sein konnten. Die Gesamtstruktur, zumal die Akzentverteilung, mag hier beigetragen haben.' That is to say, the utterance was still at least minimally distinctive ('understandable') even though the 'phoneme-features' as such were not 'present in the sound-waves'.

In the following section, a procedure is proposed for defining the phoneme as an abstractional fiction. It is suggested that the phoneme so defined is to be regarded as a heuristic or pragmatic fiction, a mere terminological convenience in describing the phonological relations which obtain among the elements of a language. Those elements are, for me, the (minimum) free forms, as defined by Bloomfield in his Language 158 ff., and in his Set of postulates.

The minimum free forms are themselves, of course, abstractions, in that each occurrence of any one of them is a fraction (approaching totality) of an acoustico-articulatory continuum which accompanies a social situation. It is a basic convention of all the social sciences that some social situations may be regarded as substantially recurrent, i.e., that certain objective differences can be disregarded. In so far as we demonstrate or assume a one-to-one correspondence between recurrent features of social situations and recurrent fractions of acoustico-articulatory continua, these latter may be regarded as linguistically recurrent. It is only a convention of linguistic science to regard these (by convention) recurrent fractions as 'elements' of a language which can be compared and from which an abstraction can be made, as a 'form'. But these conventionally or assumptionally determined abstractions appear to be as close to linguistic reality as we can at present arrive[5]; operations with them as basic elements appear to involve a minimum of undemonstrable assumption.

---

[4] Upper limit of 870 cycles.

[5] Cf. Bloomfield, Language 21 ff.

V

DEFINITION OF THE PHONEME AS AN ABSTRACTIONAL FICTITIOUS UNIT

In addition to those of the preceding paragraph, the following are the assumptions and conventions which are, I believe, necessary for this procedure of defining the phoneme. The basic assumption of all human study: Within limits knowledge is possible and communicable, and our material is within those limits. The basic concessive assumption of science: No two things and no two events[1] are exactly alike. The basic convention of science: In terms of definite criteria, if an overwhelming majority of observers agree in recognizing qualitative differentiations, the phenomena observed are different; if not, the phenomena observed are the same, in terms of those criteria. A corollary, the basic convention of linguistics: Utterances are acoustico-articulatory events accompanying social situations and correlated to them. In terms of that accompaniment and correlation, some utterances are alike and some are different in a given community.[2] We thus set up as a criterion the observable responses of members of the community as hearers. By defining 'phonetic' as subsuming all observable phenomena associated with an utterance, we deduce that, as events, all utterances must be phonetically different. But in a given community, in terms of hearers' responses in recurrent social situations, all utterances which are the same are phonetically significantly alike; and all utterances which are different are phonetically significantly different. Laboratory phonetics furnishes a generalization developed inductively from an adequate body of qualitative pointer-readings: Corresponding to significant phonetic differences are specifically articulatory differences.

With these conventions established, we can, I believe, proceed to define an abstractional fiction and apply to it a name which might as well be 'phoneme' as anything else. The procedure is here given, first in consecutive development, later expanded and illustrated.

---

[1] By regarding 'things' as entities of more than molecular magnitude, and 'events' as changes in position of such things, we avoid discussions of the identical nature of such abstractions as atoms and quanta.

[2] Cf. Bloomfield, Language 78, 144.

1. In a given community there are like utterances and different utterances. Utterances are phonetic events.

2. The forms abstracted from utterances which are the same are the same forms. Forms which are the same correspond to phonetically significantly like events. Such forms are called phonologically alike.

3. Some of the forms abstracted from utterances which are different are different forms. Excepting homonyms, forms which are different correspond to phonetically significantly different events. Such forms are called phonologically different.

4. Corresponding to phonologically different forms are significantly different phonetic events. Each of these phonetic events is composed of phonetic fractions.

5. A minimum phonological difference (greater than zero) among forms corresponds to a consistently recurring difference of one fraction of each of the phonetic events which correspond to the forms.

6. A group of forms which are minimally phonologically different constitutes an ordered class of forms.

7. This group is determined as a class by the fractional similarities of the phonetic events which correspond to the several forms.

8. This group is an ordered class, its ordering being determined by the nature of the consistently recurring phonetic differences among the fractions of the phonetic events which correspond to the several forms.

9. The relations among the members of a class of phonologically minimally different forms are minimum phonological differences. The term of any minimum phonological difference among forms is called a MICRO-PHONEME.

10. Corresponding to every phonetic difference, there is a specifically articulatory difference. That articulatory difference can be observed and recorded as a consistently recurring qualitative difference among successive phonetic events which correspond to phonologically different forms.

11. In a class of minimally phonologically different forms, the ordering (see paragraph 8) corresponds to the qualitative articulatory differences of the corresponding phonetic events.

12. Two or more such classes of minimally phonologically different forms are similarly ordered, if the qualitative articulatory differences among the corresponding phonetic events are similar and in a one-to-one relation.

13. In articulatorily similarly ordered classes of forms, the differences (i.e. ordering) of the several classes are respectively similar. The

micro-phonemes, as terms of these similar and one-to-one differences, are similarly ordered in classes. The sum of all similarly ordered terms (micro-phonemes) of similar minimum phonological differences among forms is called a MACRO-PHONEME.

Before an expanded and illustrated exposition of this technique of defining the phoneme is in order, it may be proper to state certain preliminary definitions.

A 'class' is used in the mathematical sense of 'group', usually a group the members of which share some common peculiarity. Thus the fingers of a hand may be considered as constituting a class, sharing as they do the common peculiarity of being part of the same hand, or attached to the same palm. A class is an 'ordered class', and the members of the class are 'ordered terms', when the class may be regarded as displaying some principle of organization within the class. The five fingers of a given hand are thus an ordered class, the organization of which is determined by (let us say) the relative spatial positions of the various fingers to each other with reference to the wrist, or to a given point on the palm. The finger-classes of two individuals are similarly ordered classes, for two unmutilated hands display similar relative spatial positions of the various fingers to each other with reference to the wrists, or to a given point on the palms. A's middle finger may, in and of itself, resemble B's ring finger as to length; but those two fingers are not similarly ordered in their respective classes. A's middle finger and B's middle finger, though perhaps objectively quite dissimilar, are similarly ordered in their respective classes; they are similarly ordered terms of similarly ordered classes.

We may set up an abstraction: the sum of all middle fingers of unmutilated human hands. Those various middle fingers will unquestionably present fairly wide variations in size, shape, and color. But despite the wide objective variations, all those middle fingers are recognized as similarly ordered terms of similarly ordered classes. We may go further, and speak of a 'set' of similarly ordered terms—say, all right middle fingers. In that case, we have a fraction of the sum-abstraction earlier set up.

Within the frame of this procedure of definition, a term of an ordered class of minimum phonological differences among forms is a micro-phoneme; the sum of all similarly ordered terms of similarly ordered classes is the abstraction called a macro-phoneme; a fraction of that sum-abstraction is a set of micro-phonemes.

(In the expanded exposition which follows, the numbering of the various sections is retained to correspond to that of the earlier condensed version.)

1. In a given community there are like utterances and different utterances. Utterances are phonetic events.

2. The forms abstracted from utterances which are the same are the same forms. Forms which are the same correspond to phonetically significantly like events. Such forms are called phonologically alike.

3. Some of the forms abstracted from utterances which are different are different forms. Excepting homonyms, forms which are different correspond to phonetically significantly different events. Such forms are called phonologically different.

It cannot be too strongly emphasized that a form is not a part of any specific utterance. An utterance is an event; a form is an abstraction from a large number of utterances. If I say 'That's a fine lamp', 'Won't you light the lamp?', 'Where's your old lamp?', each of these acts of speech constitutes an utterance. As a human being, I make utterances and respond to them. As a physicist, I may record and study, as physical events, the sound-waves which occur. As a physiologist, I may observe, record, and study the muscular movements which occur. As a student of language, I abstract from utterance-events certain fractions which I assume to be substantially recurrent, and (in the example above) I call that abstraction the form *lamp*. The utterance occurs, it is speech, 'parole'; the form exists, so to say, it is a part of language, 'langue'.

Throughout the rest of this discussion, these two levels must be borne in mind. We must operate with the abstracted forms and their relations; but these forms are not themselves susceptible of observation. Accordingly we observe the utterance-fractions which correspond to a form; we study the abstracted form which corresponds to utterance-fractions. By way of emphasizing the distinction between forms and utterance-fractions, the form (abstraction) is italicized; the utterance-fraction (event) is put in quotation marks. This distinction is further reinforced by the use of 'phonetic' to characterize the relations of events, of 'phonological' to characterize the relations of the abstracted forms. The inconvenience of operating on these two levels must be conceded; but the distinctions involved are indispensable.

The form *lamp* is an abstraction of a large number of utterance-fractions; it is all that is similar and nothing that is different in those utterance-fractions. When we speak of the form *lamp*, we ignore some objective differences between the corresponding fractions of various utterances, such as the differences between an act of speech by a feeble old woman and by a vigorous young man.

Forms, therefore, are either alike or different; they are also phonologically alike or phonologically different. The events which correspond to forms are all phonetically different (for no two events are exactly alike). The events which correspond to the 'same' form are phonetically significantly alike, for by 'significant' we indicate the correlation of an utterance to a social situation. The phonetic events 'Light the lamp' as produced by two different individuals are objectively very different; in so far as those events evoke similar responses in similar social situations, the two events are phonetically significantly alike. In speaking of utterances as phonetic events, we use the terms 'phonetically significantly alike or different'. In speaking of the forms which correspond to events, the terms 'phonologically alike or different' suffice. For the form is all that is similar and nothing that is different in the various events; thus only significant differences are phonologically relevant.

If two forms are phonologically different, there must be a consistently recurring phonetic significant difference between the two sets of events from which the forms are abstracted. If we say that *lamp* and *limp* are phonologically different forms, we must assert (and should demonstrate) that in any pair of phonetic events corresponding to these forms, a phonetic significant difference recurs consistently. Although A's 'lamp's are not like B's 'lamp's, nor A's 'limp's like B's 'limp's, if we can find that in A's utterances his 'limp's are characterized by a more forward and higher tongue position during some intervals than his 'lamp's, and similarly in B's utterances, then the difference recurs consistently, and there is a corresponding phonological difference between the forms which are abstracted from both speakers' utterances. The various objective variations in A's successive 'limp's or in B's successive 'limp's do not recur consistently, and therefore do not correspond to any differentiation of the form *limp*. In so far as A's 'limp's differ consistently from B's 'limp's, that difference is a personal peculiarity, of no linguistic relevance. It is only the significant differences between 'limp' and 'lamp' in everybody's utterances that concern us here; and corresponding to those consistently recurring differences of everybody's utterances is the phonological difference between the abstracted forms *limp* and *lamp*.

4. Corresponding to phonologically different forms are significantly different phonetic events. Each of these phonetic events is composed

of phonetic fractions. These fractions are usually reckoned in terms of intervals during which the articulatory organs are substantially at rest, or are leaving such a position; intervals during which some articulatory organs are making a unidirectional movement; or intervals during which some articulatory organs are moving toward a position, maintain it substantially, and leave it. These fractions are, in short, what we call 'speech-sounds', i.e. time-intervals of the actual utterance-events. In theory, these intervals might be reckoned in terms of either acoustic event (sound-waves) or articulatory event (behavior of the vocal organ muscles); in practice, the considerations of greater accessibility and susceptibility of analysis dictate the choice of the articulatory procedure.

5. A minimum phonological difference (greater than zero) among forms corresponds to a consistently recurring difference of one fraction of each of the phonetic events which correspond to the forms. A phonological difference among forms has been defined as corresponding to a consistently recurring phonetic difference among fractions of utterance-events. If we have properly reckoned the phonetic fractions, the minimum difference *which can recur consistently* as between two phonetically significantly different events is a difference of one of the fractions of each event. Corresponding to such a consistently recurring difference of one fraction of the various phonetic events is a minimum phonological difference among the forms which are abstracted from the phonetic events. Thus, in American English, the forms *beet* : *bit* : *bait* : *bet* : *bat* are minimally phonologically different. If we consider any pair of these forms, we find that a pair of the corresponding phonetic events will be characterized by a consistently recurring phonetic difference of their second fractions.

6. A group of forms which are mimimally phonologically different constitutes an ordered class of forms. Before we can conveniently operate with the group of forms *beet, bit, bait, bet, bat,* we must indicate wherein the group constitutes a class, and wherein the terms of that class are ordered.

7. This group is determined as a class by the fractional similarities of the phonetic events which correspond to the several forms. In the group of forms *beet, bit, bait, bet, bat,* the significant similarity of the beginnings and ends of the corresponding phonetic events corresponds to a phonological sameness which can be represented by the symbols [b] and [t]. [b] corresponds to all that which is similar in the beginnings of the events, i.e. closure of the lips with simultaneous vibration of the vocal cords. And [t] corresponds to all that which is similar in the ends

of the events, i.e. presence of the tongue tip and blade in contact with the alveolar ridge with simultaneous non-vibration of the vocal cords. We do not assume that the beginnings of all the events are absolutely identical; we merely assert that objective observation of the events reveals some similarity. Corresponding to that similarity, to whatever degree it can be determined, is a phonological sameness, which we indicate with the symbol [b]. We dare not operate with this phonological sameness as anything positive; all that we can assert is that these forms *beet, bit, bait, bet, bat* are not phonologically different as to their beginnings and endings. We discover from observation that the corresponding events display some objective similarity; we discover from experiment and observation that within the range of this similarity (voiced bilabial stop, voiceless alveolar stop), no differences in response can be correlated to differences in the initial and final fractions of these utterance-events. We can thus assert that our group is a phonological class of forms: *beet* : *bit* : *bait* : *bet* : *bat*, the class [b–t].

8. This group is an ordered class, its ordering being determined by the nature of the consistently recurring phonetic differences among the fractions of phonetic events which correspond to the several forms. The forms which are members of the class *beet* : *bit* : *bait* : *bet* : *bat* are abstractions from phonetic events. Any pair or any class of forms are phonologically different if there is a consistently recurring phonetic difference among the events which correspond to the forms. The phonological relations among forms thus correspond to the consistently recurrent phonetic differences among utterance-fractions. The nature of the consistently recurring phonetic differences is determined by observation. In so far as this observation can be tabulated and classified, the nature of the phonetic differences can be determined. The phonological differences of the abstracted forms are thereby correspondingly determined. In our class, we may indicate the terms of the phonological differences by the symbols [i, ɪ, e, ϵ, æ]. The difference between any pair of symbols indicates the phonological difference between the forms involved; and that phonological difference corresponds to a consistently recurring phonetic difference of the utterance-fractions. Thus, the relation [i : æ] indicates a phonological difference which corresponds to the consistently recurrent difference between A's 'beet's, B's 'beet's, C's 'beet's, ... N's 'beet's as compared with A's 'bat's, B's 'bat's, C's 'bat's, ... N's 'bat's. Those phonetic differences can be determined, and the nature of the various phonetic differences in any speaker's 'beet : bit : bait : bet : bat' can be classified; those phonetic differences

constitute an ordered class of relations. Correspondingly, the phonological differences abstracted from the utterance-fractions uttered by A, B, C, . . . N constitute an ordered class of relations; and the various terms of those relations are the terms of minimum phonological differences.

9. The relations among the members of a class of phonologically minimally different forms are minimum phonological differences. The term of any minimum phonological difference among forms is called a MICRO-PHONEME.

The differences obtaining in the series of micro-phonemes [i : ɪ : e : ɛ : æ] are abstracted from the phonological differences of the forms *beet* : *bit* : *bait* : *bet* : *bat*. This series of forms constitutes a phonologically ordered class, in American English. It is determined as a class by the constant phonological sameness of the beginnings and ends of the forms, [b–t]. Its ordering is determined by the nature of the consistently recurring phonetic differences of the corresponding phonetic events 'beet' : 'bit' : 'bait' : 'bet' : 'bat'. We indicate the abstractional phonological differences which correspond to these consistently recurring phonetic differences with the symbolical relations [i : ɪ : e : ɛ : æ]. The terms of these phonological minimum differences are micro-phonemes, as compared with each other and with reference to the class. Thus the [i] of the form *beet* is not a micro-phoneme per se; it is a micro-phoneme only in comparison with the [ɪ] of the form *bit*, etc. Further, neither of these micro-phonemes can thus far be brought into any relation with any portion or characteristic of such forms as *seek*, *sick*, *pier*. The micro-phonemes we have so far determined are the terms of the abstracted phonological differences of the forms of the class [b–t]. The [i] of *beet*, as a micro-phoneme, corresponds to everything in all the utterances 'beet' which is consistently and recurringly different from the utterances 'bit', 'bait', etc. If the nature of the labial closures in utterances 'beet' and 'bat' differs consistently, then that difference is included in the corresponding phonological difference which is represented symbolically as [i : æ].

10. Corresponding to every phonetic difference, there is a specifically articulatory difference. That articulatory difference can be observed and recorded as a consistently recurring qualitative difference among successive phonetic events which correspond to phonologically different forms.

From the preceding developments, we can make the following summaries: A phonological difference among forms corresponds to a consistently recurring phonetic difference among utterance-events. A minimum phonological difference corresponds to a difference of one phonetic fraction. A micro-phoneme corresponds to a term of a consistently recurring phonetic difference in one fraction of various utterance-events.

Laboratory phonetics furnishes us with the generalization that corresponding to every phonetic difference there is a specifically articulatory difference. The facts of articulation are accessible to observation; we can observe and tabulate the consistently recurring articulatory differences which correspond to phonetic differences. We may therefore set up a new and more practicable set of correspondences: a phonological difference corresponds to a consistently recurring articulatory difference among utterance-events. A micro-phoneme corresponds to a term of a consistently recurring articulatory difference in one fraction of various utterance-events.

11. In a class of minimally phonologically different forms, the ordering (see paragraph 8) corresponds to the qualitative articulatory differences of the corresponding phonetic events.

Let us consider, for example, the classes of forms: I. *pill* : *till* : *kill* : *bill*; II. *fin* : *thin* : *sin* : *shin*; III. *nap* : *gnat* : *knack* : *nab*. On the basis of articulatory observation of utterance-events which correspond to these forms, we may describe the consistently recurring articulatory differences among the events. We may represent the various component terms of those articulatory differences as follows: 1 bilabial; 2 aspirated; 3 voiceless; 4 exploded stop; 5 alveolar; 6 palato-velar; 7 voiced; 8 unaspirated; 9 unexploded stop; 10 labio-dental; 11 interdental; 12 dental-alveolar; 13 alveolar-palatal; 14 fricative; 15 slit narrowing; 16 groove narrowing.[3] Then we may represent our classes as being articulatorily ordered as follows:

I		II		III	
pill	1–2–3–4	*fin*	10–3–14–15	*nap*	1–8–3–9
till	5–2–3–4	*thin*	11–3–14–15	*gnat*	5–8–3–9
kill	6–2–3–4	*sin*	12–3–14–16	*knack*	6–8–3–9
bill	1–8–7–4	*shin*	13–3–14–16	*nab*	1–8–7–9

It is obvious that the same procedure is valid if the final stops of phonetic events corresponding to the forms of class III are exploded; the choice of the unexploded form is merely in the interests of reducing the similarity and rendering the connection less immediately apparent.

---

[3] This list of certain articulatory components is very sketchy, of course, omitting as it does factors of muscular tension and the time-factors of duration and acceleration. The list suffices, though, for our present purposes.

It is understood that these descriptions have to do only with relations. For laboratory phonetics does not tell us that the initial fraction of each event corresponding to the form *pill* is an absolutely identical articulatory complex. Laboratory phonetics merely tells us that as between the initial fractions of events corresponding to *pill*, *till*, etc., a consistently recurring articulatory difference has been observed, and that that difference is therefore a qualitatively constant differential relation. When we write '*pill* 1–2–3–4 (bilabial aspirated voiceless exploded stop)', we indicate that the initial fraction of any event 'pill' is marked by closure of the lips as compared with any 'till'; by a following aspiration and by openness of the glottis as compared with any 'bill'; by an exploded stop mode of articulation as compared with any 'fill'. Since the forms in each class are minimally phonologically different, the consistently recurring articulatory differences of the corresponding phonetic events determine the ordering of the class of forms.

The articulatory differences must be indicated separately from the articulatory similarities within each class. We may indicate the relations among our forms thus:

I	II	III
pill (1–2–3)—(4)	*fin* (10–15)—(3–14)	*nap* (1–3)—(8–9)
till (5–2–3)—(4)	*thin* (11–15)—(3–14)	*gnat* (5–3)—(8–9)
kill (6–2–3)—(4)	*sin* (12–16)—(3–14)	*knack* (6–3)—(8–9)
bill (1–8–7)—(4)	*shin* (13–16)—(3–14)	*nab* (1–7)—(8–9)

Thus the events 'pill' are differentially related to the events 'till' and events 'kill' in that they do not begin with an alveolar or palato-velar fraction; the events 'pill' are differentially related to the events 'bill' in that they do not begin with an unaspirated or voiced fraction.

12. Two or more such classes of minimally phonologically different forms are similarly ordered, if the qualitative articulatory differences among the corresponding phonetic events are similar and in a one-to-one relation.

If we compare our classes I and III, we find that they are articulatorily similarly ordered. The consistently recurring articulatory differences of the corresponding utterance-events are similar differences for the two classes, and in a one-to-one relation. In class I, there is a form, *pill*, corresponding to events which begin with a fraction that is not alveolar, not palato-velar, not voiced. In class III, there is a form, *nap*, corresponding to events which end with a fraction that is not alveolar, not palato-velar, not voiced. Similarly for each term of the two classes. The relations within the two classes are similar, in that the relations (1–3) : (5–3) : (6–3) : (1–7) recur in both classes. The relation 2 : 2 : 2 : 8, which we find in class I, must be left out of account; for the factor (unaspirated) 8 is one of the constant articulatory factors of class III. We conclude that the differentiation of aspirated and unaspirated exploded stops in phonetic events is correlated to certain sequences within the event (initial 'position' before stressed vowels), and that in these sequences the differentiation of aspirated and unaspirated stops is parallel to the differentiation of voiceless and voiced stops.

Thus although there is no similarity between classes I and II or classes II and III, between classes I and III there is a similarity of ordering. These classes are in a one-to-one relation: Given the order *pill—till—kill—bill* or any other order of class I, there is one and only one order of class III *nap, gnat, knack, nab* which satisfies the requirements of a one-to-one relation. This order must be found by trial and error, since we make no assumptions concerning phoneme-membership. (The arrangements I have given are for purposes of exposition adapted to reveal that relation at once; that is, so to say, accidental.) Thus we may say that our classes

pill	(1–3)—(2)	*nap*	(1–3)
till	(5–3)—(2)	*gnat*	(5–3)
kill	(6–3)—(2)	*knack*	(6–3)
bill	(1–7)—(8)	*nab*	(1–7)

are similarly ordered, for the phonological differences among the forms of the two classes correspond to a consistently recurring articulatory set of differences among the phonetic events, and those articulatory differences are similar and one-to-one in the two classes, as correlated with the classes of forms.

Since the classes I and III of forms are phonologically similarly ordered, the phonological differences among the terms of classes I and III are similarly ordered. Above (paragraph 9) we have defined a micro-phoneme as the term of a minimum phonological difference. Since the differences of classes I and III are similarly ordered, the micro-phonemes which are terms of these differences are similarly ordered.

13. In articulatorily similarly ordered classes of forms, the differences (i.e. ordering) of the several classes are respectively similar. The micro-phonemes, as terms of these similar and one-to-one differences, are

similarly ordered in the classes. The sum of all similarly ordered terms (micro-phonemes) of similar minimum phonological differences among forms is called a MACRO-PHONEME.

Let us consider the micro-phonemes which are terms of the phonological differences of classes I and III.

pill. . . . (1–3) . . . . *nap*
till . . . . (5–3) . . . . *gnat*
kill. . . . (6–3) . . . . *knack*
bill . . . . (1–7) . . . . *nab*

Putting it diagrammatically: We have hitherto seen in these forms only a vertical relation. We have noted differences; these differences have been determined as phonological differences. Later we have seen that these differences can be correlated to consistently recurring articulatory differences of the corresponding events. But now we have established that these differences (the orderings of the two classes) are similar and in a one-to-one relation. Now we can begin to establish horizontal relations; we may connect with a dotted horizontal line the similarly ordered terms (micro-phonemes) of the similarly ordered (vertical) classes. Thus we set up a new abstraction, and say that the micro-phoneme which is correlated to *pill* belongs with the micro-phoneme which is correlated to *nap*. We can do this, not because the articulatory complexes which correspond to the micro-phonemes are inherently identical or positively similar, but solely because those articulatory complexes are similarly differentiated from other articulatory complexes. The initial fraction of an event 'pill' is not identical with the final fraction of an event 'nap'. But the initial fraction of a 'pill' differs from the initial fraction of a 'till' in a relation which is similar to the relation wherein the final fraction of a 'nap' differs from the final fraction of a 'gnat'. (The analogy of the middle fingers will be recalled.)

The 'p-phoneme' is therefore the sum of *all* those phonological differentiae which correspond to a bilabial articulation as opposed to alveolar or palato-velar, a voiceless articulation as opposed to voiced, a stop articulation as opposed to fricative. Thus we can combine the stops of 'pill, nap, tapper', etc., as opposed to the stops of 'till, gnat, tatter', etc.; 'kill, knack, tacker', etc.; 'bill, nab, tabber', etc. Any articulatory complex which is a bilabial voiceless stop corresponds to a constituent member of the 'p-phoneme', if it occurs in a phonetic sequence where it may contrast with an articulation which is bilabial, stop, but not voiceless; (labial), voiceless, but not stop; voiceless, stop, but not bilabial.

We have no right (and we don't want) to establish any 'phonemic' relations between the stops of a 'pill' and of a 'spill', for the stop of a 'spill' is not in a phonetic sequence where it may contrast with a not-voiceless articulation. Any constituent member of the 'p-phoneme' must correspond to an articulatory complex which is in a phonetic sequence where it is different from a complex corresponding to a constituent member of the 'b-phoneme'. The stop of a 'pill' and the stop of a 'spill' cannot correspond to the same macro-phoneme, for the corresponding micro-phonemes are not similarly ordered. There is a phonological difference between the forms *pill* and *bill*; there is no similar phonological difference between the forms *spill* and **sbill.* There appears to be no alternative to considering the stops of 'spill, spare, spin', etc. as corresponding to a different phoneme from the stops of 'pill, pare, nap, lip, tapper, slapper', etc. The stops of 'spill, spare', etc. are significantly bilabial and stop, but not significantly voiceless; the stops of 'pill, nap, tapper', etc. are significantly bilabial, stop, *and voiceless.*

When we speak, then, of a (macro-)phoneme, we are using an abstraction as a terminological convenience to describe the recurrence of similar phonological differentiations among the elements of a language. It cannot be too strongly emphasized that the phoneme, so defined, is meaningless as applied to any particular linguistic element: it is a negative, relational, differential abstraction; it is a unit of that sort of relation which de Saussure describes[4]: 'Dans la langue il n'y a que des différences sans termes positifs.' It follows, therefore, that it is meaningless to speak of 'the third phoneme (micro- or macro-) of the form *sudden*', or to speak of 'an occurrence of a phoneme'. What occurs is not a phoneme, for the phoneme is defined as the term of a recurrent differential relation. What occurs is a phonetic fraction or a differentiated articulatory complex correlated to a micro-phoneme. A phoneme, accordingly, does not occur; it 'exists' in the somewhat peculiar sense of existence that a brother, *qua* brother, 'exists'—as a term of a relation. Linguists who accept as important de Saussure's distinction between 'parole' and 'langue' may find it helpful to regard the differentiated articulatory complex as a fraction of 'la parole', the phoneme as a term of those relations which comprise the system of 'la langue'; of course we can admit no mentalistic assumptions as to the real existence of the phoneme in the souls or what not of the speakers of that 'langue'.

There is no denying that this definition is complicated and forbidding.

---

[4] Cours de Linguistique Générale (1922) 166.

But the definition of any fundamental unit is always difficult, and must be particularly difficult in the case of linguistic phenomena, where we are dealing with events in time and bodies in space, with recurrence and uniqueness of events. The only defense that may be offered for this procedure is that it appears to be characterized by a minimum of the undemonstrable. With one coherent set of assumptions and conventions, which are indispensable to all scientific linguistic study, and one sound laboratory generalization, we may apply strictly mathematical methods and deduce a logically unimpeachable definition of some entity, which we may call a (macro-)phoneme. Any errors in reasoning are accidental to the author's own deficiencies and are fortunately detectible; as between a sound and a merely simple procedure there can be no choice.

Two corollary details of procedure should be mentioned.

Few, if any, languages present what we may call an absolutely symmetrical recurrence of phonological differences, i.e. it is not always possible to match classes of micro-phonemes in a one-to-one relation. We can set up in American English five phonemes, corresponding to the front vowels [i, ɪ, e, ɛ, æ][5] in the series *beet, bit, bait, bet, bat*; *seek, sick, sake, sec, sack*; etc. But in the frame [f–t] the series is incomplete: *feet, fit, fate, ——, fat.* The question is whether the absence of a form [*fɛt] is a fortuitous or fundamental characteristic of the [f–t] series. To decide such questions, we may admit a procedural convention: We find, in American English, that all fundamental characteristics involving the absence of (presumably potential) distinctive forms can be correlated with immediately preceding or following phonetic fractions, including the omnipresent factor of stress. We have already seen, from such examples as *bet, met, set*, etc., that we can discover phonemic relations involving the fourth term of the series [i : ɪ : e : ɛ : æ] before voiceless alveolar stops. We test for relations with preceding voiceless labio-dental fricatives, and find such series as *fane* : *fen* : *fan*; *fade* : *fed* : *fad*; *phase* : *fez*; etc. We conclude that the absence of the phonetic fractions [fɛ-] and [-ɛt] is not a fundamental characteristic of

---

[5] Here, as elsewhere, I make traditional use of the IPA symbols in transcribing American English rather than Bloomfield's, not because I regard that traditional use as inherently better—for I consider Bloomfield's transcriptions of themselves simpler and more rational—but quite simply because phonetic transcriptions, for the reasons given elsewhere in this study, are arbitrary and conventional; they partake at least in part of the nature of orthographies, and accordingly an established tradition of transcription is preferable to any new one which is based on the same principles. Cf. R.-M. S. Heffner, Concerning Transcription, in LANGUAGE 10.288 ff.

American English; we conclude that the absence of a fourth term in the series of forms *feet* : *fit* : *fate* : — : *fat* is fortuitous. The discovery of *fetlock* is an unnecessary support of our conclusion.

This convention must be applied with caution. For the language of some Americans, it appears to be fortuitous, for others fundamental, that the front-vowel series is always defective before voiced palatovelars in monosyllables. In cases of doubt, the burden of proof is of course upon the one who regards such an absence as fortuitous; for the language itself, as it actually is, is our final authority.

The stops of 'writing, latter' furnish another kind of difficulty, which leads to almost ridiculous inconsistencies in the attempt to apply some definitions of the phoneme. These stops, in many dialects of General American English, are definitely voiced; yet the corresponding events are distinctively different from 'riding, ladder'. The micro-phonemes corresponding to *latter* and *ladder* must be correlated to some articulatory difference; that difference may be formulated in terms of duration: the first stop is definitely shorter than the second. We can correlate these two stops with those of 'tin' and 'din', if the first of these is also regularly shorter—as it appears to be from the tables presented by Fletcher.[6] In the case of 'tin : din', the difference in duration of the initial phonetic fractions is by no means the most conspicuous or 'distinctive' difference. But if that difference is regularly present here, and is present (and 'distinctive') in 'latter : ladder', we may regard the two classes of micro-phonemes as similarly ordered with respect to an articulatory factor, and combine the micro-phonemes corresponding to *tin*, *latter* in one macro-phoneme, those corresponding to *din*, *ladder* in another. This combination is of course not justified by any positive similarities between the stops, but only by the similarity of the relations of one factor of their articulations. If there were no such forms as *riding*, *ladder*, the stops of 'writing, latter' could no more be correlated with that of 'tin' than can the stop of 'still'. Cases of this sort, where the important articulatory difference corresponding to one class of micro-phonemes is similar to a relatively minor, apparently insignificant articulatory difference corresponding to another class of micro-phonemes, are fairly rare; but any definition of the phoneme must be prepared to meet them.

It goes without saying that an analogous procedure is to be followed in defining phrase and sentence phonemes, corresponding to stress, pitch, tempo, etc., in so far as they constitute phonetic differentiae among utterances.

[6] Speech and Hearing 60.

VI

COMPARISONS WITH EARLIER DEFINITIONS OF THE PHONEME

Any comparison of this definition with the earlier ones reveals a difference in procedure, to be sure; it is also clear that the object of this definition is a new entity. Since this new entity may, however, be made use of in much the same way as what others have defined as a phoneme, I have taken the liberty of applying that name to the entity I define. The term 'macro-phoneme' is of course available to anyone who protests against that application.

The greatest difference between the macro-phoneme and the various 'phonemes' is that the former is an abstractional fiction, the latter are realities of some sort. The macro-phoneme is a term of recurrent relations; the 'phonemes' are a repertory of positive additive entities. The macro-phoneme is an abstraction from forms; the 'phonemes' are the constituent elements of forms, of phonetic events, or sometimes of both. It suffices to quote two definers of radically different 'phonemes': Bloomfield[1]: 'This word consists of such and such of the basic phonemes of my language'; and Trubetzkoy[2]; 'Aber ebenso wird dasselbe Wort *duby* als aus den Phonemen $d + u + b + y$ bestehend wahrgenommen.' The 'phonemes' for these linguists are obviously positive additive entities; in some cases, these entities are regarded as possessing inherent potencies of utilizable distinctiveness, or as offering to the speaker a repertory of autonomous code-signals. Thus Daniel Jones speaks[3] of 'the fact that separate phonemes are capable of distinguishing words'; and Jespersen refers[4] to phonemes as 'sounds that can be used for distinctive purposes, i.e. to keep words apart which would otherwise be identical'. It is plain from Jespersen's statement, quoted earlier, on the relations of phoneme and form, that this characterization of the phoneme is a much abbreviated and inverted statement of his real interpretation. But such formulations, if not significant, are at all events dangerous, as leading all too readily to a kind of mythology in which the hypostasized 'phonemes' play their rôles, or an equally mythological view of the linguistic process according to which a speaker reaches into his store of phonemes, selects the proper number of each, arranges them tastefully, and then produces an utterance.

---

[1] Modern Philology 25.217.

[2] TCLP 1.39.

[3] TCLP 4.77.

[4] Linguistica 391.

The macro-phoneme is a fiction, defined for the purpose of describing conveniently the phonological relations among the elements of a language, its forms. The sum of such relations among the elements is the phonological system of the language. This phonological system is of course nothing objectively existent; it is not definable as a mental pattern in the minds of the speakers of the language; it is not even a 'platonic idea' which the language actualizes. It is quite simply the sum of all the phonological relations among the forms of a language, as those relations are determined by objective study. The phonological system is the phonetician's and phonologist's summarized formulation of the relations; it is not a phenomenon, nor an intuition.

There are various kinds of phonological relations in a given language. These relations may be classified into groups, corresponding to the differentiations in a given articulatory range. When we say that American English, as a phonological system, has a maximum of six front vowel and diphthong phonemes, we mean that there are, within the range of that mode of articulation which we call vowel or diphthong, at most fifteen[5] significant phonetic differentiations of utterances by differentiation in the positions and movements of the front of the tongue. By taking all the articulatory ranges within which significant differentiation occurs in the language, we can construct the maximal phonological system of the language. But in nearly all languages, that potential maximum of differentiation is not uniformly realized. There are systematic gaps, i.e. there are cases of less than maximal differentiation within a given articulatory range. These systematic gaps are just as much a part of the phonological system as are the maximal differentiations. To illustrate with the vowels and diphthongs of my dialect of American English:

---

[5] $(6-1) + (6-2) + (6-3) + (6-4) + (6-5)$. Thus, *bite* : *beet, bit, bait, bet, bat*; *beet* : *bit, bait, bet, bat*; *bit* : *bait, bet, bat*; *bait* : *bet, bat*; *bet* : *bat*. Some readers may prefer the formulation: If x is the maximum number of significant phonological differentiations within a given articulatory range in a language, then $2x = n(n-1)$, where n is the maximum number of phonemes in that range, and $(n-1)$ is the number of consecutive phonological relations within that range.

bite	beet	bit	bait	bet	bat	bot	bought	boat
like	leak	lick	lake	——	lack	lock	Lawk	——
pyre	pier		pair			par	pour	

———	boot	bout	butte	Bert	butt
look	Luke	——	——	lurk	luck
poor		power	pure	purr	

It is clear that the differentiations of vowels and diphthongs before [-r] are just as systematic and just as 'symmetrical' as the quite dissimilar differentiations before [-t, -k]. There is *a* vowel system in monosyllables before [-r] in American English, but it is not the same vowel system as before [-t, -k]. And to call either of these systems *the* American English vowel system is to suppress an important fact about the phonological system of the language. To force the phonemes corresponding to the vowels of our third row into some violent and arbitrary classification in terms of the first two rows is almost (if not quite) as stultifying as to try to fit French vowel-phonemes into an English system. The phonological relations which obtain in such a series of forms as *pyre* : *pier* : *pair*, etc., are not similar nor comparable to the relations in such a series as *bite* : *beet* : *bit* : *bait*, etc. Since a phoneme has been defined as the sum of terms of recurrent phonological relations, it follows that the phonemes which correspond to the two series are not comparable. The phonemes which correspond to the [-r]-series must be different phonemes from those corresponding to the [-t]-series. Similarly, the phoneme corresponding to the stop in 'spill' is a different phoneme from the one corresponding to the stops in 'pill, tapper, nap'.

The relatively large number of such phonemes in a given language will doubtless appear unfortunate to some linguists. Most of their objections, I believe, can be met if the purpose of the phoneme as a terminological unit is considered. For me, the phoneme is a unit defined for a convenient description of phonological relations. Naturally, then, I need different phonemes to describe different phonological relations. For many linguists, it appears, the phoneme functions as a unit to be represented by a symbol in so-called phonetic transcriptions. It appears that the unit these linguists require cannot sufficiently take into account either phonological or phonetic facts; it would clarify the issue if these units might be called 'graphemes', 'transcribemes', or even 'letters'.

For I know of no earlier phoneme-definition which does not achieve transcriptional sanctions by violence to essential phonological relations and palpable phonetic fact.

It follows that, according to this definition, there are as many different phonemes in a language as there are consecutive relations of significant phonetic differentiation (plus one) in each articulatory range. That is, if we have in American English five consecutive relations in *bite* : *beet* : *bit* : *bait* : *bet* : *bat*, there are six phonemes for these and similar relations; if we have two consecutive relations in *pyre* : *pier* : *pair*, there are three phonemes for these and similar relations; polysyllables with accented vowel before [-r-] will display another set of relations, which may not correspond to either of the earlier series.

The number of phonemes in a given language is thus limited; that number is a sub-multiple of the number of forms of the language. Therein one must agree with those definitions of the phoneme which specify that the phonemes of a language are strictly limited in number. But that number is frequently set much lower than this definition will admit; that number is usually set in terms solely of the maximum differentiations within all articulatory ranges. (The coincidence of this circumstance with the fact that such a number of written symbols suffices for transcription purposes—if some arbitrary conventions are adopted—need not surprise us.) But if we want to define some unit for purposes of describing the phonological relations within a language, we must reconcile ourselves to having as many of these units as there are recurrent phonological relations within the language.

Jespersen, in his Monosyllabism in English[6], estimates, on the basis of permutations and combinations of the maximal variations within all articulatory ranges, that the number of possible English monosyllables is 'rather more than 158,000'. The exact number does not concern us; it is enough to know its general order, and to know that the actual number of English monosyllables represents but a fraction of it. In what might be called 'the potential monosyllabic vocabulary' of English, there are thus numerous gaps. Some of those gaps are unquestionably 'fortuitous' (cf. [*fɛt] above); but equally unquestionably some of those gaps are systematic. In so far as the latter is the case, those gaps are a part of the phonological system of English; the systematic absence of those 'potential monosyllables' is a highly important phonological relation within the language.

---

[6] Linguistica 390.

VII

ARTICULATORY RATHER THAN ACOUSTIC CHARACTERIZATION OF THE PHONEME

The choice of articulatory rather than acoustic fractions as the basis for classifying the phonetic relations will scarcely need justification. It is generally assumed that there is a correlation between the articulation and the sound-waves. That correlation is further assumed to be in a sense absolute, i.e. an observable difference in articulation is correlated with *some* observable difference in the sound-waves. But the articulation is a three-dimensional complex, which may further be practically analyzed into components on the basis of muscular movements; the sound-waves are a uni-dimensional event (for purposes of analysis), and the analysis, as a Fourier series, etc., leads to components which are not readily comparable with the articulatory components. When Trubetzkoy asserts[1]: 'Phonetisch betrachtet ist ein Vokal eine Verbindung mehrerer Artikulationsakte, von denen jeder eine bestimmte akustische Wirkung besitzt', if that assertion implies that acoustic analysis and articulatory analysis lead to a one-to-one correspondence of components, it is an unverified and very improbable assumption. Jespersen's Wesen und System der Sprachlaute[2] may be pertinently adduced here, especially the section: 'Ebenso ist es wenigstens höchst zweifelhaft, ob wir berechtigt sind, z.b. ein gewöhnliches stimmhaftes [o] als aus der durch geflüstertes [o] hervorgerufenen mundresonanz + stimmton bestehend zu betrachten. Und doch haben wir keinen zweifel inbetreff des engsten zusammenhangs der laute: ein gesungenes [o], ein gesprochenes [o], ein geflüstertes [o], ein gehauchtes [o] und ein geblasenes [o] fühlen wir als denselben vokal in verschiedenen zustandsformen, die nicht einmal so verschieden sind wie eis, wasser und dampf, weil man schon unmittelbar die einheit wahrnimmt. Es muss, wie es scheint, hier etwas gemeinschaftliches vorhanden sein, das die wiedererkennung bedingt; das gemeinsame im gebiete Mund ist die stellung aller organe oberhalb der stimmbänder; aber was ist das gemeinsame in den gebieten Luftwellen und Ohr? Das hat uns kein akustiker sagen können, und—es *gibt vielleicht* gar nichts gemeinsames weder in den schwingungen der luft noch im eindruck des ohres. . . . ' Until the successful refutation of this line of argument (which applies a fortiori to the stops of *pit* and *tip*), the study of phonetic differences in terms of articulation is the indicated procedure.

---

[1] TCLP 1.41.

[2] Linguistica 134 ff.; quotation from 143 f.

The determination of the objective articulatory facts of the mode of differentiation in connection with classes of micro-phonemes is the study of speech-sounds, a study which must precede the determination of the phonemes of a language, and which profitably may be carried on after and in terms of that determination. Any discovery of objective similarity in the articulatory complexes which correspond to similar terms of micro-phoneme classes is a confirmation of the correctness of the groupings. But that similarity is not an absolute one, hence the impropriety of using it as the basis for the groupings. We combine the stops of 'pin, sopping, nap' as corresponding to a single phoneme, not because of any constant quantitative characteristic, but because of a constant qualitative relation to the stops of 'bin, sobbing, nab' (among other relations).

The discovery of certain similarities[3] among the articulatory complexes, which correspond to similar terms of classes of micro-phonemes, is not surprising. Indeed, that similarity is a basic assumption (usually tacit) of most phonetic study and even of many definitions of the phoneme. For these classes of articulatory complexes represent a distribution of differentiation within a given articulatory range. Given a one-to-one correspondence of distributed differentiation within a constant range, there is a high degree of probability of an approximation of the differentiated objective phenomena. The nature of the approximation is of course for the psychologist to study, in terms of imitation and habit. But it is not certain that the factors of imitation and habit are the sole determinants of the objective articulatory phenomena. The distribution of differentiation within a given articulatory range may be a factor in itself. There is of course no automatic adjustment of the mode of articulation to an optimum of communication-efficiency; nor is there any conscious effort on the part of speakers of a language toward such an adjustment. But the possibility of long-range partial adjustment as an important factor in phonetic change is apparent upon realistic consideration of the way in which children are trained to speak, i.e., to produce utterances which are sufficiently differentiated, and differentiated in articulatory modes which are sufficiently similar to those prevailing in their community.

---

[3] That is, not only are the classes *beet* : *bit* : *bait* : *bet* : *bat* and *seek* : *sick* : *sake* : *sec* : *sack* similarly differentiated, but there is considerable objective similarity between the vowels of 'beet' and 'seek', 'bit' and 'sick', etc.

VIII

SOME POSSIBLE APPLICATIONS OF AN ABSTRACTIONAL FICTITIOUS PHONEME

It should be clear for what kinds of procedure in linguistic study a unit like the (macro-)phoneme is adapted. It is a procedure very much like the 'phonology' of the Cercle linguistique de Prague. That such a procedure is in order can scarcely be questioned; there will be few students of language who disagree seriously with Trubetzkoy's statement[1]: 'Eine vergleichende Untersuchung der phonologischen Systeme aller Sprachen der Erde ist die dringendste Aufgabe der heutigen Sprachwissenschaft.' The only limitation which the definition proposed above would impose upon such a procedure is the necessity for antecedent and concurrent phonetic (and articulatory) analysis. If the valuable and suggestive work of many members of the Cercle linguistique de Prague has not been wholly convincing to many students of language, it is (aside from its newness) because of the subjective mentalistic definition of units and a somewhat truculent denial of the relevance of phonetic analysis.[2]

Many students of language will agree, I believe, that the tasks of the phonetician and of the phonologist are not antithetical or unrelated, but that phonological synthesis is an integral continuation of phonetic analysis. The phonetician, in so far as he is anything more than a laboratory technician, must also be a phonologist; the phonologist, in so far as he is anything more than an artist, must be a phonetician. It is the phonetician's failure to finish his job which the phonologist rightly protests against, not the part of the job he has done. In so far as a specialization of labor is necessary, it is for the phonetician to consider utterances as articulatory events, objectively measurable (and correlated to habits); it is for the phonologist to consider abstractions from utterances as terms of recurrent differentiations (and correlated to social situations).

The necessity for phonetic analysis as a basis for the determination of phonemes is clear. The following remarks venture to suggest some possible values of a continuation of phonetic analysis even after the phonemes have been determined.

---

[1] TCLP 1.41.

[2] For an excellent discussion of this point, see K. Bühler, Phonetik und Phonologie (TCLP 4.22 ff.); see also Jespersen, Linguistica 213.

We have seen that the relations between pairs of articulatory complexes, correlated to pairs of micro-phonemic sets[3], though qualitatively constant (if the several micro-phonemes belong to the same macro-phonemes), may be objectively different. Thus, there is a constant qualitative relation between the stops of 'tin, hat, latter' and 'din, had, ladder'. But the articulatory complexes which are the terms of those differences are not identical: there is even a quantitative variation in the degree of difference. There are more habitual articulatory differences between the stops of 'tin' and 'din' than between those of 'hat' and 'had', or of 'latter' and 'ladder'. We are justified in speaking of the position of maximal articulatory differentiation between phonetic fractions correlated to phonemes; we may even in some cases arrange the articulatory differentiations in a descending scale of degree of differentiation: thus *tin* : *din*, *hat* : *had*, *latter* : *ladder*. The scale, in this case, reaches zero degree of differentiation with such a form as *still*, where there is no differentiation (and no micro-phonemic class) within the articulatory range of alveolar stop: *still* is different only from *spill* and *skill*. This is not to say that we can 'explain' the peculiarly ambiguous (in terms of a [t-d] opposition) articulation of the stop in 'still'; it is merely a rather interesting correlation. It is possible that similar correlations might be established for the degree of difference between articulatory complexes which occur in other positions, as corresponding to sets of micro-phonemes. The determination of the 'functional burdening' which would serve as the correlative to such articulatory differentiation is however not so feasible as some operations with the term might indicate. To be sure, we can readily compute the number of pairs of forms minimally different by a given pair of micro-phonemic sets; we can even approximately compute the relative frequencies of each of these forms. But the semantic utilization of these forms in relation to one another, i.e. the possibilities of ambiguity in otherwise similar contexts, is quite impossible to measure. Yet the first two factors, or even the first alone, may be regarded with all due reserves as a rough index of 'functional burdening', if we explicitly state that phonemes are meaningless, and that any selection of forms on the basis of phonemes is therefore random sampling in terms of meaning.[4] If we can find that the 'functional burdening', as roughly indicated, is greater between [t- : d-] than between [-t : -d] or [-t- : -d-], we may set up a correlation which is of considerable importance for the study of American English, and, as a tentative principle, for the study of the linguistic process in general. And in the classifying of such correlations, if they can be found, it is possible that a specifically relational unit may prove a useful terminological convenience.

---

[3] Hereafter, the term 'set of micro-phonemes' means 'set of micro-phonemes which are similarly ordered in sequentially similar classes'. Thus the vowels of 'beet : bit' correspond to a class of micro-phonemes; the vowels of 'beet, feet' and of 'beet, bean' correspond to sets of micro-phonemes. A set of micro-phonemes is accordingly a fraction of a macro-phoneme.

The proposed terminology may also be useful in formulating diachronic relations of phonological significance. Among the events which we class as 'phonetic changes' there is a wide range of importance from the phonological point of view.[5] The results of phonetic change may or may not be a modification of the qualitative differential relations of forms in the language. If a modification of the habits of articulation occurs in all phonetic fractions corresponding to forms which stand in a certain phonemic relation to the other forms of a language, then we have a case of so-called 'absolute phonetic change'. If such a modification occurs only in phonetic fractions corresponding to forms, the relations of which constitute relations of sets of micro-phonemes (sub-classes of a macro-phonemic relation), then it is a 'conditional phonetic change'. When an absolute or phonemically total phonetic change is accompanied by similarly total changes in an entire articulatory class, the product is a phonetic shift, which, despite the extent of modification of articulatory habits, does not modify the phonological system of the language, as a system of qualitative relations among forms. The first Germanic consonant shift (with the exception of **sp*, **st*, **sk*, **pt*, **kt*, which may not be exceptions at all) is an example. Any pairs of forms which were qualitatively differentiated before such a shift are also qualitatively differentiated afterwards. But if the results of a phonetic change or changes are an alteration of the qualitative differentiations among forms, that is a fact of prime phonological importance. That between Middle High German and Modern German there has been a change which we describe as [ɛɪ > ae] is a phonetic fact; similarly the change [i: > ae]. But the combination of these phonetic facts is a phonological fact, for the relations of such forms as *ein* : *mîn* are now qualitatively modified:

---

[4] It is worth noting that this assumption of the meaningless phoneme, which alone validates the claim of random sampling and thereby justifies operation with 'functional burdening', has sometimes been explicitly denied by the very linguists who speak of 'functional burdening'.

[5] Cf. B. Trnka, Some Remarks on the Phonological Structure of English, in Xenia Pragensia 357 ff.

ein : *mein*. The results of conditional phonetic change obviously are very often of phonological importance, since they are a new differentiation or a new distribution of differentiations among the forms.

In so far as the denial of autonomy and linguistic relevance to the 'speech-sounds' as such involves a repudiation of Junggrammatiker assumptions, it is of course necessary to state a definition of phonetic change in terms of the proposed definition, so as to justify the continued use of the products of the Junggrammatiker procedure. We may define a phonetic change as a (gradual) substitution of new habits of producing differentiated articulatory complexes, correlated to sets of micro-phonemes. Then we can claim all the validity of phonetic law to which the Junggrammatiker procedure can pretend. If the sets of micro-phonemes involved are the totality of a macro-phoneme, the phonetic change is absolute. If the sets of micro-phonemes are less than totality of a macro-phoneme, the phonetic change is conditional; and the conditions are no harder to specify than under the Junggrammatiker procedure. From the point of view of the end-product, a phonetic change is a modification either of the qualitative differentiation or of the mode of qualitative differentiation among the forms of a language.

Just as above it was suggested that it might be possible to find a correlation between 'functional burdening' and the degree of articulatory differentiation, so here, in the field of diachronic study, it may be suggested that there is a possibility of correlation between a kind of 'functional burdening' and apparently conditional phonetic change. The somewhat peculiar history of the linguistic and articulatory habits associated with IE **sp*, **st*, **sk* (reduplication, throughout the various consonant shifts, in Old Germanic metrics, down to the present) may be correlated with the absence (as we conclude for other reasons) of any **sb*, **sd*, **sg* with which they could enter into phonemic relations.

The material (though not the explanations) which the 'functionalists' like Horn adduce may be susceptible of study from this point of view. An example from American English may serve: In many dialects, the preterit tenses of certain weak verbs are not regularly marked by a final dental stop. Thus [krip : krɛp; kip : kɛp; slip : slɛp; wip : wɛp]. Other weak verbs, however, always have the final dental in the preterit: [hip : hipt; ræp : ræpt; hɛlp : hɛlpt]. To be sure, the collocation of these preterits effectively destroys any fable of an exceptionless phonetic 'treatment' of [t] after [p]. And any attempt to explain the divergence on the basis of an analogy will involve some rather spectacular arguments. But if we consider the earlier forms [krip : *krɛpt—hip : hipt], it appears that in the second instance we have a differentiation

[t] : zero which corresponds to a micro-phonemic class; in the first instance, that objectively similar phonetic differentiation can be correlated only to a partially micro-phonemic class. It is not for us to say that the vowel alternation is more significant than the suffix for preterit marking of the verbs *creep, keep, sleep, weep,* or any such psychologistic formulation; it is enough to point out that the final alveolar stops in 'crept, kept, slept, wept' were not correlated to the same set of micro-phonemes as the final stops of 'heaped, wrapped, helped'. This method of classification would not conflict with the definition of a phonetic change given above, and of course does not involve the introduction of mentalistic factors.[6]

The only use of the various other 'phonemes' to which the proposed macro-phoneme cannot be put is the easy establishing of sanctions for broad transcriptions. For obviously the proposed definition justifies the use of the same symbol for the initial and final stops of 'top-hat' as correlated to the same phoneme, but not for the vowels of 'pair' and either 'pain' or 'pen'. It is difficult to feel too gloomy about this deficiency. Many students of language, probably, have felt acutely or obscurely that our conventional transcriptions are merely a somewhat less awkward but still conventionalized orthography, that phonetic transcription says either too much or too little. According to the proposed definition four kinds of quasi-phonetic transcription would be possible: (1) A more or less mathematical indication of the relations of micro-phonemes in terms of intra-class differentiations; (2) A more or less clinical indication of articulatory movements and positions; (3) A more or less subjective indication of the acoustic effect upon the transcriber as a kind of imperfect phonograph; (4) A more or less conventionalized indication of highly abstracted units of combined phonemes—which is about what our present transcriptions are—but in terms of clearly defined conventions. It would be eminently desirable to have a transcriber explicitly and expressly state that he intends to indicate the vowel of 'pair' as [e, ej, ϵ, ϵɪ, ϵj] etc., rather than imply tacitly that it is related to the same phoneme as the vowel of either 'pain' or 'pen', which is to conceal rather than reveal the phonological and phonetic facts. If he chooses to use an entirely different symbol, no one can object. At all events, a temporary embarassment in the codifying of transcription conventions may be a small price to pay if some dangerous assumptions and inconsistencies can be eliminated and clarity can be brought to the study of the phonological relations within languages.

---

[6] Cf. Modern Philology 25.225.